

UNIVERSITY OF BELGRADE
FACULTY OF ORGANIZATIONAL SCIENCES

Nenad R. Gligorić

**Recognition of students' activity during
the lecture utilizing the Internet of
Things**

Doctoral dissertation

Belgrade, June 2014.

Mentor:

docent dr Srđan Krčo

Docent

Univerzitet u Beogradu

Fakultet organizacionih nauka

Članovi komisije:

Prof. dr Božidar Radenković

Redovni profesor

Univerzitet u Beogradu

Fakultet organizacionih nauka

Prof. dr Milorad Stanojević

Redovni profesor

Univerzitet u Beogradu

Saobraćajni fakultet

Datum odbrane:

Praćenje aktivnosti studenata tokom nastave primenom Interneta inteligentnih uređaja

REZIME

Predmet ovog istraživanja je praćenje aktivnosti studenata, odnosno prepoznavanje paterna iz okruženja i njihova prezentacija posredstvom tehnologije Interneta inteligentnih uređaja. Glavna hipoteza od koje se polazi i koja je dokazana u okviru ove doktorske disertacije je da se primenom tehnologije Interneta inteligentnih uređaja u nastavi može poboljšati efikasnost nastavnog procesa kroz realizaciju sistema za praćenje aktivnosti studenata koji u gotovo realnom vremenu omogućava analizu parametara iz okruženja i prezentaciju obrađenih rezultata. Kako bi se sagledao broj tehnika neophodnih za realizaciju pomenutih procesa kao i da bi se opravdala potreba date studije, urađen je pregled relevantnih istraživanja na polju računarski i društvenih nauka. Pregledom je obuhvaćeno poređenje i komparativna analiza platformi pametnih učionica koje pre svega predstavljaju platforme u kojima će zaživeti nova sveobuhvatna kompjuterska rešenja sposobna da prepoznaju sociološke kontekste u momentu pojavljivanja. Društvene nauke su vrlo bitne da bi se razumela sama pozadina procesa koji se klasifikuju i prate, tako da je pregledom društvenih nauka, pre svega socioloških signala, zaokružen pregled relevantnih bibliografskih izvora i data smernica dalje moguće realizacije sistema za praćenje aktivnosti. Pregledom utvrđeni su potencijalni parametri i algoritmi koje je moguće analizirati i upotrebiti za analizu a potom je predstavljena metodologija koja je korišćena u istraživanju za sve faze studije. Pre definisanja algoritama koji se mogu koristiti i postavke arhitekture objašnjeni su zahtevi sistema koje je neophodno ispuniti da bi prepoznavanje paterna u učionici moglo efikasno obavljati. Prepoznavanje paterna se obavlja metodom mašinskog učenja za koji je preduslov postojanje klasifikatora baziranom na određenom setu podataka. Simulacija sistema je urađena pre implementacije korišćenjem seta podataka koja nisu korišćeni u procesu treniranja. Pri simulaciji sistem je pokazao prosečnu tačnost od 92.2%. Nakon simulacije, sistem je implementiran i perforamanse su evaluirane upoređivanjem

anotatora u realnom vremenu (ocena studenata za vreme predavanja) i ocene koje je dao sistem. Prosečna tačnost sistema evaluiranog na ovaj način za tri grupe studenata bila je 81.9%. Sistem je implementiran u Matlab-u i u mogućnosti je da ekstraktuje parametre iz digitalanog signala koji se prikupljaju pomoću kamere i mikrofona. Iz slike se ekstraktuje intezitet pokreta studenata na osnovu razlika između frejmova dok se iz zvuka ekstraktuju sledeće karakteristike: glasovni segmenti, vreme govora, glasovna stopa, broj kratkih govornih segmenata, formantna frekvencija, pouzdanost formantne frekvencije, vrednost najvećeg autokorelacionog pika, lokacija najvećeg autokorelacionog pika, broj autokorelacionih pikova i govorna aktivnost. Vrednosti ovih parametara se zatim koriste kao input za klasifikator koji se sastoji od skupa podataka ekstraktovanog iz velikog broja snimljenih predavanja, za vreme kojih su studenti ocenjivali kvalitet predavanja u realnom vremenu korišćenjem aplikacije na mobilnom telefonu. Finalna evaluacija performansi sistema sprovedena nakon implementacije, gde je sistem korišćen za prepoznavanje parametara uživo na predavanjima, pokazala je da je sistem sa zadovoljavajućom tačnošću (81.9%) u mogućnosti da prepozna zainteresovanost studenata.

Ključne reči: Internet inteligentnih uređaja; obrada digitalnog signala; klasifikacija; pametne učionice; komunikacija između mašina; sociološki signali.

Naučna oblast: Informacioni sistemi i tehnologije

Uža naučna oblast: Pervazivno računarstvo

UDK broj:

Recognition of students' activity during the lecture utilizing the Internet of Things

ABSTRACT

This paper proposes novel method for detecting students' attention by utilizing Internet of Things and machine learning algorithms. The main hypothesis that has been proven in this PhD thesis is that utilization of the Internet of Things in the education can increase teaching efficiency by implementing system for detecting students' attention that enables environmental parameters analysis and presentation of processed results. In order to provide further insight required for realization of above mentioned processes as well as to justify the need for performing featuring study, a survey on relevant computer and social science researches is done. The survey includes comparative analysis of smart classroom platforms that represents a medium for these new algorithms to rise and identify sociological contexts in a moment of appearance. The sociological sciences are very important as they help us understand a background of sociological processes being monitored and classified; thus survey of sociological sciences and social signals above all are given to finalize a survey of relevant researches and inline a direction of further realization of the proposed system. In addition, survey has also inlined potential parameters and algorithms that can be used, followed with a methodology description for all phases of the research. Then, requirements of such system are analyzed and important features required for detection identified as well as sociological factors that influence these features. Pattern classification is done by leveraging a machine learning method that requires classifier based on a certain dataset. Before implementation, system simulation is done on a dataset which is not used in the process of training. During simulation system have shown average accuracy of 92.2%. After the simulation, the system was implemented and its performance evaluated by comparing a real-time annotator (i.e. the students' feedback) with the system output during the lectures. The average accuracy of the system evaluated for three different groups of students was 81.9%.

Keywords: Internet of Things; Digital Signal Processing; Classification; Smart Classrooms; M2M communication; Social Signals.

Scientific field: Information systems and technology

Scientific subfield: Pervasive computing

UDK:

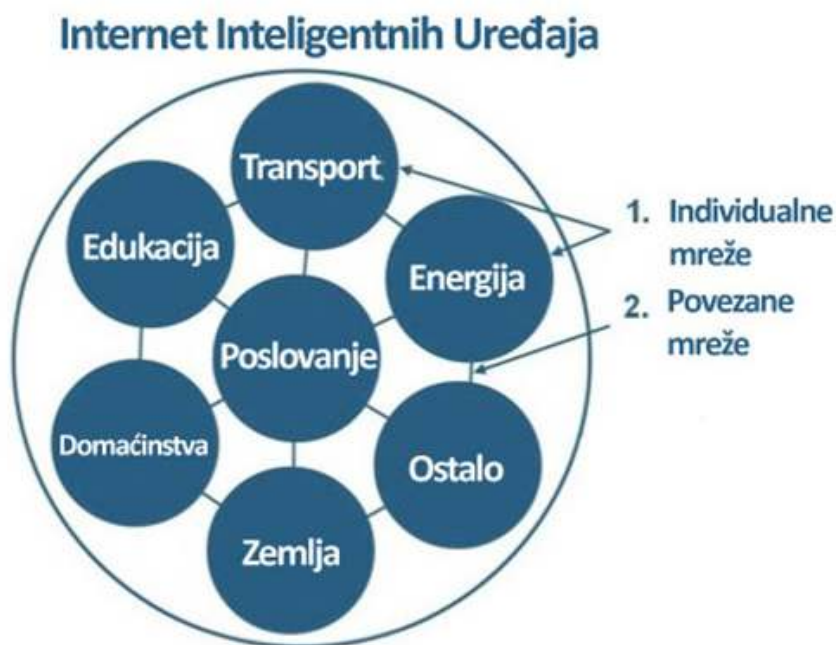
Sadržaj

1. Uvod.....	1
1.1. Definicija problema i predmet istraživanja.....	3
1.2. Ciljevi istraživanja u doktorskoj disertaciji.....	5
1.3. Polazne hipoteze.....	6
1.4. Metode istraživanja.....	7
1.5. Očekivani doprinos rada.....	8
1.6. Plan istraživanja i struktura rada.....	10
2. Pregled ostvarenih rezultata i relevantnih istraživanja.....	13
2.1. Koncept pametne učionice.....	15
2.1.1. Pametne učionice, pojam i definicija.....	15
2.1.2. Kategorizacija pametnih učionica.....	16
2.2. Komparivni pregled pametnih učionica.....	18
2.3. Platforme inteligentnih okruženja.....	29
2.4. Uloga društvenih nauka.....	29
2.4.1. Sociološki signali.....	29
2.4.2. Analiza socioloških signala.....	30
2.4.3. Ekstrakcija socioloških signala.....	35
2.4.4. Metode segmentacije i klasifikacije socioloških signala.....	43
2.4.5. Psihologija.....	44
2.4.5. Ostala istraživanja.....	45
2.5. Potencijalni parametri.....	46
2.5.1. Ostala istraživanja.....	47
2.6. Zaključak.....	67
3. Metodologija istraživanja.....	70
3.2. Ekstrakcija parametara.....	70
3.2.1. Ekstrakcija parametara iz zvuka.....	71
3.2.2. Ekstrakcija parametara sa slike.....	75
3.2.3. Merenje provodnosti kože.....	78
3.3. Analiza i statistička obrada podataka.....	80
3.4. Metodologija evaluacije performansi sistema.....	81
4. Zahtevi sistema za praćenja aktivnosti studenata.....	84
4.1. Sistemski zahtevi.....	85
4.1.1. Razlike između trenažnog i izvršnog okruženja.....	85
4.1.2. Performanse sistema.....	85
4.1.3. Simultano izvršavanje radnih procesa.....	86
4.1.4. Hardverski sistemski zahtevi.....	86
4.2. Integritet sociološke interakcije.....	86
4.2.1. Mesto postavljanja IoT uređaja.....	86
4.3. Skladištenje i protok informacija.....	87

4.4. Otvorena arhitektura.....	87
4.5. Presentacija podatka.....	88
5. Obzervaciono istraživanje i prikupljanje podataka za klasifikaciju.....	89
5.1. Anketiranje uzorka.....	89
5.1.1. Rezultati ankete.....	89
5.2. Posmatranje uzorka i izgradnja dataset-a.....	90
5.2.1. Anotacija podataka.....	91
5.2.3. Rezultujući set podataka.....	95
6. Analiza parametara.....	98
6.1. Parametri iz zvučnog signala.....	98
6.2. Parametri ekstraktovani sa slike.....	104
6.3. Galvanska provodnost kože.....	105
6.4. Korelacija parametara.....	108
7. Evaluacija mašinskih metoda učenja.....	110
7.1. Izbor klasifikatora.....	110
7.2. Selekcija parametara.....	114
8. Dizajn i implementacija sistema za praćenje aktivnosti studenata.....	117
8.1. Modelovanje sistema za praćenja aktivnosti studenata.....	117
8.1.1. Opis sistema.....	117
8.1.2. Arhitektura sistema za praćenje aktivnosti studenata.....	117
8.1.3. Dijagram slučajeva korišćenja.....	118
8.1.4. Sistemski dijagram sekvenci.....	121
8.1.5. UML dijagram klasa.....	123
8.2. Implementacija sistema za praćenje aktivnosti studenata.....	124
8.2.1. Sloj prikupljanja podataka.....	124
8.2.2. Sloj za obradu i analizu signala.....	127
8.2.3. Sloj za prezentaciju i skladištenje.....	132
9. Evaluacija performansi sistema u realnom okruženju.....	138
9.1. Analiza performansi klasifikatora i rezultata istraživanja.....	140
10. Zaključak.....	143
10.1. Kritički osvrt na sprovedeno istraživanje.....	144
10.2. Budući pravci istraživanja.....	145
6. Literatura.....	146
Biografija.....	162
Izjava o autorstvu.....	165
Izjava o istovetnosti štampane i elektronske verzije doktorskog rada.....	166
Izjava o korišćenju.....	167

1. Uvod

Krajem prošlog veka prvi put se pominje nov koncept objedinjenog prisustva širokog spektra uređaja povezanih na Internet - Internet inteligentnih uređaja (eng. Internet of Things - IoT). IoT omogućava nove vidove komunikacije između ljudi i objekata kao i samih objekata (eng. Machine to Machine – M2M) sa tendencijom da promeni perspektivu informacionog društva. Fenomen Interneta inteligentnih uređaja zahteva jedinstvene šema adresiranja i komunikacije što je prouzrokovalo nastanak novih protokola i potrebu dodatne standardizacije. IoT ima potencijal da promeni sve aspekte ljudskog života i smatra se da će biti jedna od deset najznačajnijih tehnologija koje će uticati na dalji tehnološki razvoj u budućnosti (Evans, 2011).



Slika 1 Internet inteligentnih uređaja

U ovom radu opisana je primena IoT-a u edukaciji kako bi se tokom nastave pratila aktivnost studenata. Dato istraživanje može se kategorizovati u oblast pametnih učionica, i efikasno se dopunjava sa tehnologijama koje se inače koriste u ovom slučaju. Pametne učionice su nastale automatizacijom ambijenta čime se nepotrebni procesi svode na minimum tako da rad i učenje postaju efikasniji. Predavanje u pametnim

učionicama se obavlja na tradicionalan način ili se koriste mešoviti tipovi učenja (eng. Blending learning) koji su nastali kao rezultat spajanja nekoliko vrsta podučavanja. Trenutno, ispitivanje kvaliteta procesa predavanja bazira se na povratnom toku informacija dobijenih sa zakašnjenjem i sadržajno nepotpunim za detaljniju analizu. U globalu, nezavisno od pristupa, evidentan je nedostatak mehanizma koji bi dozvolio studentima da izraze nejasnoće, ili predavaču da stekne uvid u stepen razumevanja predavanja od strane studenata. IoT se može u konceptu pametne učionice upotrebiti za razvoj dinamičkog automatizovanog sistema za praćanje pažnje u realnom vremenu. Aktivnost studenata je kompleksna pojava koja se identifikuje preko skupa parametara i kao takva zahteva razumevanje širokog spektra biheviorističkih i socioloških elemenata koji je definišu. Prema tome, jedan deo rada predstavlja analizu sociološkog fenomena u procesu edukacije s ciljem da se opravdaju metode izabrane za praćenje aktivnosti studenata.

Značaj IoT-a je u velikom broju podataka koje je moguće prikupiti o jednom entitetu zahvaljujući IoT uređajima koji se primenjuju u različitim kontekstima i njihovoj proaktivnosti u realnom vremenu baziranoj na predefinisanim akcijama. Primena IoT-a u edukaciji je višestruka i prevazilazi praćenje aktivnosti studenata; bilo koji uređaj koji se koristi u procesu edukacije se može priključiti na Internet gde je broj scenarija dalje upotrebe neograničen. U slučaju praćenja aktivnosti studenata informacije prikupljene na ovaj način značajne su i za dalja istraživanja u oblasti računarskih i društvenih nauka.

Pravac svih informacionih tehnologija ide u smeru veštačke inteligencije, a tehnički zahtevi i metode koje se koriste pri realizaciji pametne učionice, široko su primenljive i van date oblasti. U bliskoj budućnosti akcenat je na senzorskim mrežama, i objedinjenoj spoznaji entiteta koji nas okružuju, upotrebom Interneta inteligentnih uređaja. Senzorski uređaji obezbeđuju podatke, koji prikupljeni u određenom kontekstu zajedno sa materijalom zabeleženim pratećom opremom, mogu doprineti razvoju interesantih platformi koje će dalje oblikovati tok savremene edukacije. Predložen sistem može postati standardna komponenta pametne učionice kako bi se omogućio uvid u nedostatke predavanja u momentu prezentacije. Ukoliko se proces edukacije u celosti oslanja na mogućnosti tehnološkog napretka, postignut rezultat neće biti dovoljno

efikasan. Naprotiv, kvalitativne metode i strategije podučavanja je potrebno koristiti uporedo sa tehnološkim prednostima koje donose pametne učionice.

1.1. Definicija problema i predmet istraživanja

Predmet istraživanja ovog rada je primena Interneta inteligentnih uređaja u obrazovanju, kako bi se pratila aktivnost studenata tokom nastave radi automatskog određivanja pažnje i stepena zainteresovanosti. Cilj istraživanja je realizacija sistema koji bi omogućio praćenje socioloških procesa posredstvom senzora i monitoring uređaja, kao i alata za analizu i klasifikaciju.

Proces predavanja predstavlja skup različitih tehnika koje se koriste pri prezentaciji, čija efikasnost primene zavisi od mnogobrojnih faktora, kako bi se prepoznala potreba promene pristupa predavanju; npr. primenio pristup sa više interakcije, ili ponovila određena deonica ili neke druge metode koja će za određeni slučaj dati efikasnije rezultate.

Kao što je već naglašeno, u ovom momentu, ispitivanje kvaliteta procesa predavanja bazira se na povratnom toku informacija dobijenih sa zakašnjenjem (upitnici popunjeni nakon predavanja) kao i sadržajno nepotpunim za detaljniju analizu (stvaranje subjektivnog mišljenja o stepenu razumevanja kroz interakciju i postavljanjem pitanja na temu predavanja, itd). U globalu nezavisno od pristupa, evidentan je nedostatak mehanizma koji bi dozvolio studentima da izraze nejasnoće, ili predavaču da stekne uvid u stepen razumevanja predavanja od strane studenata. IoT se može u konceptu pametne učionice upotrebiti za razvoj dinamičkog automatizovanog sistema za praćenje pažnje kako bi se prepoznavanjem određenih paterna iz auditorijuma doneo zaključak o kvalitetu predavanja i same satisfakcije studenata istim.

Istraživanja na temu implementacije inteligentnih servisa praćenja aktivnosti u okruženju tradicionalne učionice do sada nije bilo, iz razloga jer se zahteva spoznaja

parametara koji pripadaju multidisciplinarnim oblastima: sociološkim naukama (biheaviorističkih procesi, sociološki signali) i računarskim naukama (tehnologije koje se koriste za identifikaciju i analizu biheaviorističkih procesa), kao i zbog složenosti socioloških procesa uopšte.

Sintezom različitih oblasti, stiču se uslovi za realizaciju sistema pametne učionice koji u gotovo realnom vremenu prati aktivnosti studenata i pruža povratnu informaciju predavaču, koji na osnovu iste, može da prilagodi pristup predavanju prema potrebama studenata. Interaktivni procesi između predavača i auditorijuma tokom nastave zahtevaju razumevanje sociološke i psihološke pozadine aktivnosti koja se prati. Analiziranje signala generalno zahteva analizu svih paterna koji okružuju vizuelnu i/ili verbalnu poruku i utiču na njeno značenje ili karakterišu neku sociološku ulogu.

U edukaciji, sociološki signali studenata mogu da se koriste kao smernice koje bi preko povratnog kanala u realnom vremenu ukazivale na nedostatke predavanja, stepen razumevanja prezentovane materije i zainteresovanost auditorjuma. U većini slučajeva verbalna komunikacija je praćena ovim signalima, koji se nazivaju i „iskreni“ signali zbog svog podsvesnog karaktera, čime se odaju informacije kao psihološka pozadina određenih postupaka. Oblast koja se bavi istraživanjem socioloških signala s ciljem da se tehnologija unapredi u prirodnijem i fleksibilnijem pravcu naziva se *Social Aware Computing*. Međutim, da bi se ostvarili ciljevi istraživanja postavljeni u ovoj studiji, neophodno je spojiti više polja istraživanja od Sve-obuhvatnog ili pervazivnog računarstva, preko Interneta inteligentnih uređaja do procesiranja socioloških sigala i društvenih nauka uopšte.

Istraživanja u ovoj disertaciji obuhvatiće: modelovanje sistema za praćenje aktivnosti studenata korišćenjem metodologija, kao i razvoj mehanizama za analizu, klasifikaciju i prezentaciju digitalnog signala u formi pogodnoj za krajnjeg korisnika. Studija slučaja u okviru disertacije ima za cilj realizaciju predloženog modela kroz implementaciju sistema za praćenje aktivnosti studenata tokom nastave primenom tehnologija Interneta inteligentnih uređaja.

Materija koja će se obrađivati u ovoj disertaciji obuhvata:

- Istraživanje postojećih rešenja inteligentnih sistema koji se koriste u obrazovanju, klasifikacija i jedinstvena kategorizacija istih
- Identifikacija parametara koji se direktno ili indirektno mogu upotrebiti za donošenje zaključka o stepenu zainteresovanosti auditorijuma. Pregled relevantnih tehnologija za obradu signala, analiza i utvrđivanje mogućnosti primene
- Analiza i razvoj metoda za identifikaciju i prikupljanje prethodno definisanih paterna i njihova klasifikacija
- Evaluaciju performansi sistema preko simulacije koja obuhvata treniranje, testiranje i validaciju klasifikacije na različitim setovima podataka prikupljenih za vreme predavanja
- Analizu dobijenih rezultata i korelacije svih parametara uključenih u proces klasifikacije i prepoznavanja parametara
- Pripremu realizacije automatizovanog sistema baziranog na Internetu inteligentnih uređaja na osnovu predloženih modela kroz metode projektovanja i modelovanja
- Realizaciju sistema upotrebom različitih softverskih i hardverskih alata; evaluacija sistema, testiranje sistema i upotreba korektivnih mera

1.2. Ciljevi istraživanja u doktorskoj disertaciji

Primarni cilj istraživanja doktorske disertacije je modelovanje i realizacija sistema za praćenje aktivnosti studenata upotrebom IoT-a, kako bi se u realnom vremenu tokom predavanja obezbedilo praćenje i uvid u stepen zainteresovanost auditorijuma.

Najvažniji cilj koji se postiže istraživanjem, jeste povećavanje efikasnosti procesa predavanja koji kao krajnji rezultat ima pozitivan uticaj na proces edukacije uopšte.

Cilj ovog istraživanja preciznije definisano obuhvata:

- Utvrđivanje mogućnosti upotrebe Interneta inteligentnih uređaja za praćenje aktivnosti studenata za vreme nastave
- Definisanje modela za praćenje aktivnosti identifikacijom parametara
- Evaluaciju kreiranog modela kroz implementaciju sistema i testiranje u celini

S obzirom na navedene ciljeve, zadaci istraživanja su:

- Analiza metoda za realizaciju sistema i sistematizacija relevantnih bibliografskih istraživanja; kao i identifikacija parametara koji se mogu iskoristiti za detektovanje paterna pažnje
- Definisanje modela za različite ulazne vrednosti parametara, njihova korelacija i klasifikacija prema datim modelima
- Analiza postojećih tehnika za obradu i analizu signala; Inicijalna merenja predloženih parametara za vreme procesa predavanja i analiza dobijenih rezultata
- Projektovanje arhitekture sistema za praćenje aktivnosti studenata tokom nastave
- Evaluacija tačnosti klasifikacije realizovanog sistema

Materija koja bi se obrađivala u radu je: Istraživanje mogućnosti implementacije sistema koji u realnom vremenu obezbeđuje praćenje aktivnosti studenata; Identifikaciju parametara koji se direktnoj ili indirektnoj vezi sa donošenjem zaključka o stepenu zainteresovanosti auditorijuma; Utvrđivanje i razvoj metoda za identifikaciju i prikupljanje prethodno definisanih paterna; Definisanje tehničkih zahteva za praćenje prepoznatih parametara; Pripremu realizacije automatizovanog sistema korišćenjem IoT-a na osnovu predložene arhitekture; Evaluacija performansi realizovanog sistema na setu podataka različitog od seta korišćenog za treniranje klasifikatora.

1.3. Polazne hipoteze

Glavna hipoteza razvijena u okviru istraživanja polazeći od postavljenih ciljeva i zadataka istraživanja glasi:

Primenom Interneta inteligentnih uređaja u nastavi može se poboljšati efikasnost nastavnog procesa kroz realizaciju sistema za praćenje aktivnosti studenata koji omogućava analizu parametara iz okruženja i prezentaciju obrađenih rezultata.

Pomoćne hipoteze:

H1. Praćenjem aktivnosti studenata tokom nastave pomoću Interneta inteligentnih uređaja mogu se postići efikasniji rezultati u odnosu na tradicionalan pristup predavanju

H2. Upotrebom tehnologije Interneta inteligentnih uređaja moguće je prepoznati sociološke procese aktivnosti studenata

H3. Primenom metoda identifikacije i klasifikacije paterna moguće je utvrditi nivo zainteresovanosti studenata i realizovati sistem za praćenje aktivnosti

- H3a. Karakteristike ekstraktovane iz zvuka i pokreta studenata mogu se upotrebiti za klasifikaciju i prepoznavanje pažnje studenata i njihove zainteresovanosti

1.4. Metode istraživanja

U disertaciji od opšte naučnih metoda koristiće se: metoda analize, analitičko-deduktivna metoda, komparativna metoda, merenje i metoda naučnog posmatranja i eksperimenta. Osnovni metod u radu je metoda analize dosadašnjih rezultata i relevantnih bibliografskih istraživanja. Komparativna metoda koristiće se prilikom upoređivanja postojećih rešenja inteligentnih sistema kao polazne osnove za nadogradnju planiranog sistema. Od opšte naučnih metoda koristiće se analitičko-deduktivna metoda prilikom analize podataka prikupljenih za vreme eksperimenata.

Od tehnika za prikupljanje podataka, koristiće se metoda naučnog posmatranja i metoda istraživanja uzorka. Metoda naučnog posmatranja primeniće se prilikom praćenja studenata za vreme predavanja i određivanja parametara koji identifikuju paterne

zainteresovanosti. Merenje relevantnih parametara i analiza dobijenih rezultata obaviće se pomoću standardnih statističkih metoda.

Naučno razvrstavanje (klasifikacija) primeniće se kod prepoznavanja paterna za vreme simulacije i finalne evaluacije sistema.

1.5. Očekivani doprinos rada

Najvažniji rezultat istraživanja u okviru ove disertacije predstavlja sistem za praćenje aktivnosti studenata tokom nastave primenom IoT-a, koji dalju upotrebnu vrednost ima u procesu edukacije kao samostalan sistem ili kao komponenta pametne učionice koja se može integrisati sa sistemom za elektronsko obrazovanje ili sistemom za učenje na daljinu. Originalnost istraživanja ogleda se u jedinstvenom metodološkom pristupu identifikacije socioloških paterna aktivnosti studenata koja se obavlja procesiranjem digitalnog zapisa, i prezentacijom rezultata u gotovo realnom vremenu kako bi se nastavni proces učinio što efikasnijim. Tehnički zahtevi i metode neophodne za realizaciju ciljeva istraživanja, uveliko su primenljivi i van date oblasti. Paterni na osnovu kojih je moguće doneti zaključak o stepenu zainteresovanosti univerzalni su za većinu socioloških konteksta. Zbog kompleksnosti i multidisciplinarnosti teme istraživanja, do sada nije bilo radova koji su bavili istraživanjem ove oblasti. Prema tome, materija obrađena ovom disertacijom predstavlja naučni doprinos i ima posebnu vrednost.

Prema tome naučni i stručni doprinosi doprinos rada su:

- Definisane modela za praćenje aktivnosti studenata tokom nastave primenom IoT i okvirnog modela za njegovo rešavanje posredstvom Interneta inteligentnih uređaja. Definisanjem metodologije omogućava se implementacija sistema za praćenje aktivnosti studenata koji može da doprinese poboljšanju procesa predavanja.

- Jedinstvena kategorizacija inteligentnih okruženja i detaljan prikaz njihove primene. Pregled postojećih platformi pametne učionice i jedinstvena kategorizacija prema tehničkim karakteristikama i funkcijama prvi put je definisana u ovom istraživanju.
- Identifikacija i analiza paterna studenata koji se mogu koristiti za određivanje stepena njihove zainteresovanosti i njihova korelacija. U ovom istraživanju, po prvi put su identifikovani paterni iz grupnog govora i pokreta studenata i data je analiza i korelacija istih.
- Evaluacija parametara upotrebom različitih statističkih metoda i finalna selekcija parametara koji se mogu koristiti za prepoznavanje aktivnosti studenata. Finalni skup parametara omogućava efikasno prepoznavanje aktivnosti studenata korišćenjem samo relevantnih parametara.
- Definisano metoda za implementaciju sistema za praćenje aktivnosti studenata upotrebom Interneta inteligentnih uređaja. Definisanjem metodologije omogućava se implementacija sistema za praćenje aktivnosti studenata koji može da doprinese poboljšanju procesa predavanja.
- Evaluaciji različitih metoda mašinskog učenja za detekciju zainteresovanosti predavanja. Evaluacija performanse devet algoritama koji se trenutno najviše upotrebljavaju pri klasifikaciji različitih konektsta i selekcija algoritma sa najboljim rezultatima među evaluiranim algoritmima.
- Razvoju sistema za praćenje aktivnosti koji će omogućiti praćenje i prikaz aktivnosti u auditorijumu, ekstrakcijom različitih paterna i njihovu analizu putem uređaja povezanih na Internet; kao i skladištenje, prezentaciju analiziranog materijala i interfejs za pristup od strane trećeg lica (eksternih platformi). Razvoj sistema proširiće perspektivu trenutnog procesa edukacije, prikazom nedostataka u momentu prezentacije materijala. Sistem je evaluiran u

realnom okruženju i rezultati su upoređeni sa rezultatima dobijenim simulacijom.

Ostali doprinosi

- Pregled oblasti i jedinstvena kategorizacija pametnih učionica koja predstavlja sistematizaciju i kompletan uvid u stepen trenutnog razvoja; ističe preplitanje i intezitet kojim se različite tehnologije dopunjuju, obrazujući kompleksne platforme nastale kao rezultat implementacije različitih tehnologija pametne učionice u sopstvenoj režiji
- Sistematizacija različitih socioloških/psiholoških istraživanja koja su u vezi sa temom rada i korelaciona sinteza sociološkog i tehološkog modela u jedinstven model

1.6. Plan istraživanja i struktura rada

Plan istraživanja je dat u tabeli 1:

Tabela 1. Dinamika istraživanja

FAZA	ZADACI	METODE, TEHNIKE, STANDARDI
1. Analiza postojećih modela i rešenja iz oblasti	<ul style="list-style-type: none"> • Prikupljanje podataka od opšteg značaja anketiranjem studenata. • Pretraživanje baza podataka. • Pretraživanje stručne 	<ul style="list-style-type: none"> – Pretraživanje baza podataka – Petraživanje stručne literature – Pretraživanje Internet

	literature. Anketa	resursa – Studije slučaja
2. Obzervaciono istraživanje	<ul style="list-style-type: none"> • Prikupljanje podataka različitim metodama za klasifikaciju 	<ul style="list-style-type: none"> – Anketa – Snimanje predavanja – Anotacija podataka
3. Predlog modela sistema za praćenja aktivnosti studenata.	<ul style="list-style-type: none"> • Sistematizacija postojećih tehnika za obradu i analizu signala • Definisane modela za različite ulazne vrednosti parametara • Klasifikacija parametara za date modele • Inicijalna merenja predloženih parametara za vreme procesa predavanja • Analiza dobijenih rezultata • Istraživanje tehnika za razvoj sistema 	<ul style="list-style-type: none"> – Pretraživanje baza podataka – Petraživanje stručne literature – Pretraživanje Internet resursa – Studije slučaja – Statistička klasifikacija – Merenja na određenom uzorku
4. Projektovanje modela i servisa	<ul style="list-style-type: none"> • Definisane zahteva, analiza podataka. • Analiza arhitekture. Projektovanje komponenta	<ul style="list-style-type: none"> – Metode projektovanja – UML

5. Razvoj sistema za praćenje aktivnosti studenata	<ul style="list-style-type: none"> • Razvoj sistema i servisa • Razvoj dodatnih servisa i aplikacija • Integracija komponenata 	<ul style="list-style-type: none"> – C, Matlab – Java, Android
5. Evaluacija performansi sistema	Analiza performansi klasifikatora i rezultata istraživanja	<ul style="list-style-type: none"> – Evaluacija performansi klasifikatora – Statistička analiza

2. Pregled ostvarenih rezultata i relevantnih istraživanja

Relevantni bibliografski izvori i ostvareni rezultati koji su u vezi sa predmetom istraživanja iz oblasti pametnih učionica, usmereni su pre na automatizaciju procesa predavanja i prateće opreme, nego na podizanje nivoa obrazovanja direktnim uticanjem na pažnju auditorijuma.

U cilju istraživanja uticaja, postojeće metode polaze od pretpostavke da se koncepti visokog nivoa potencijalno mogu razlikovati od obzervacije signala na drugim nivoima i da se ti modeli mogu identifikovati. Basu et al. (Cooper et al. 1956) opisuje metodu za automatsko detektovanje uticaja pri interakciji u prostoriji gde su učesnici u posmatranju igrali interaktivne igre. Uticaj je modelovan upotrebom dinamičkih Bajesovih mreža koji identifikuje grupnu interakciju kao grupu markovljevih lanaca, od kojih svaki utiče na promenu stanja drugog.

U jednom drugom istraživanju (Di et al. 2008) predloženo je korišćenje mašinskog učenja pod nadzorom baziranog na percepciji problema kao troklasnoj klasifikaciji, iz uzorka snimljenog za vreme sastanaka i upotrebom metode potpornih vektora.

U studiji (Madam et al. 2004), zainteresovanost grupe je određena upotrebom IoT uređaja merenjem karakteristika iz audio signala, pokreta glave i galvanske provodnosti kože, nakon čega je kombinacija ovih parametara upotrebljena da bi se dobio grupni indeks zainteresovanosti. Individualne i grupne govorne karakteristike su ekstraktovane korišćenjem više-nivoskog Markovljevog modela i uparenog skrivenog Markovljevog modela. Potvrдно klimanje glavom, koje se u mnogim kulturama koristi kao znak potvrđivanja, mereno je uporebom mešovitog Gausovog modela (DeVaul, et al. 2003).

Metoda za analizu zvuka iz pomenute studije od koristi je za ovo istraživanje. Kontekst kao i način prikupljanja parametara i analiza istih se razlikuju u ovom istraživanju. Analiza u ovom radu uključuje i pokrete studenata. Zvuk u studiji (Madam et al. 2004) je prikupljen korišćenjem mikrofona postavljenih u blizini govornika, nasuprot

integrisanom mikrofONU kao sastavom delu video opreme korišćene u ovom istraživanju. Zainteresovanost grupe u studiji (Madam et al. 2004) je utvrđena analizom promene vrednosti provodnosti kože (eng. Galvanic Skin Response), koja predstavlja vezu između događaja u okruženju i psihološkog stanja pojedinca (Cooper et al. 1956); dok je u ovom istraživanju to urađeno analizom parametara iz zvuka i pokreta i njihovom korelacijom; a galvanska provodnost kože merena je samo kod jednog studenta radi utvrđivanja verodostojnosti ocene predavanja koju je student dao za vreme obzervacione faze istraživanja.

Do sada je bilo nekoliko istraživanja koja su koristila GSR da bi se odredilo afektivno stanje pojedinca. StratleCam (Healey & Picard, 1998) sistem je mobilna platforma koju korisnik nosi sa sobom i koja meri provodnost kože i koristi tu informaciju da u momentima zainteresovanosti korisnika sačuvava slike iz njegovog okruženja. Pored toga što je StratleCam mobilni sistem, za vreme eksperimenata korisnik je bio statičan i izložen audio stimulansima.

Provodna jakna (Nakra, 2000) koristi GSR senzor u kombinaciji sa drugim sensorima koje mere psihološke karakteristike, da bi se napravila korelacija provodnosti sa emocijama u muzičkim sekvencama.

U svakom od istraživanja, istaknut je značaj pokreta, čiji intenzitet bi trebalo da bude što manji da bi tačnost GSR senzora bila što veća. Uticaj pokreta na GSR, predmet je studija (Healey, 2000) (Westeyny et al. 2006) gde su subjekti izvršavali akcije sedenja, šetanja, trčanja i kašljanja, čime je dokazana korelacija između merenja koja su pokazala različiti senzori.

U daljem tekstu pregledom su obuhvaćeni svi sistemi inteligentnih okruženja koji se koriste u obrazovanju: opisane su postojeće platforme i po prvi put data je njihova jedinstvena kategorizacija i napravljen je pregled relevantnih istraživanja iz oblasti. Kao

osnova daljeg razvoja, pojavljuje se potreba razumevanja socioloških i psiholoških aspekata praćenih aktivnosti, konkretno socioloških signala.

Posle pametnih učionica, definisani su sociološki signali, prikazan je pregled tehnika koje obezbeđuju njihovo prikupljanje i analizu kako bi se ukazala potreba razvoja novih i prikazao skup postojećih tehnologija koje mogu izaći u susret aktivnom praćenju aktivnosti studenata u realnom vremenu.

2.1. Koncept pametne učionice

Pametne učionice mogu da koriste senzore za prepoznavanje pokreta ili glasa korisnika radi ekstrakcije određenog paterna na osnovu koga se izvršava predefinisana akcija. Automatizacijom ambijenta, ostaje više vremena za predavanje, time što se vreme ne troši na nepotrebne procese pripreme. Predavači koji ranije nisu radili u ovakvom okruženju, susreće se sa otežavajućim okolnostima, dok ne savladaju osnove korišćenja.

2.1.1. Pametne učionice, pojam i definicija

Pametna učionica (eng. Smart Classroom) je nastala po uzoru na pametne prostore (Shi et al, 2002) (Pishva, 2007), a njen funkcionalan okvir se može definisati kao inteligentno okruženje opremljeno različitim spektrom hardverske i softverske opreme: projektorima, kamerama, sensorima, modulima za prepoznavanje govora, lica, itd (Xie et al. 2001). U literaturi se ovaj termin pojavljuje i pod nazivom inteligentna učionica.

Širenjem funkcionalnosti, povećava se broj aplikacija i komponenti koje se koriste: od samog početka inteligentni prostori prate tehnološki razvoj primenom poslednjih naučnih dostignuća. Prema tome, pametna učionica je nova sinteza tehnologije, korisničkog interfejsa i tradicionalnih metoda predavanja radi stvaranja inovativnog, naprednog i fleksibilnog okruženja za učenje (Stiefelbogen et al. 2006).

U praktičnoj primeni postoji mnogo različitih implementacija pametne učionice, npr. učionice bazirane na sistemima za video konferenciju, rešenja za učenje na daljinu bazirana na multimedijalnom okruženju, učionice koje podržavaju video i čet, učionice sa virtuelnim 3D prikazom korisničkih avatara koje imaju mogućnost audio i video interakcije, itd (Xin et al. 2010).

Primetan je nedostatak podele na osnovu koje se može izvršiti kategorizacija različitog skupa komponenti i funkcija svih pametnih učionica. Radi lakšeg razgraničenja tehnike i aplikacione pripadnosti različitih implementacija, po prvi put data je jedinstvena podela pametnih učionica.

2.1.2. Kategorizacija pametnih učionica

Komponente pametnih učionica prema funkcionalnoj operativnosti mogu se podeliti na sledeće kategorije (Gligorić et al. 2011):

- virtuelni asistenti
- automatizovano snimanje predavanja
- digitalizacija pisanih materijala
- sistemi za video/audio konferencije
- virtuelizacija učesnika

Virtuelni asistenti (Xu et al, 2009) mapiraju ljudski govor i pokrete sa računarskim akcijama koristeći ih kao instrukcije, npr. menjanje slajda na glasovnu komandu ili paljenje projektora na detektovano logovanje korisnika. U ovu grupu se mogu svrstati i uređaji koji funkcionalno zamenjuju više drugih uređaja; npr. laserske olovke koje se koriste umesto miša, markera, i sl.

Automatizovano snimanje predavanja je kategorija sistema koji omogućavaju automatsko snimanje predavanja, praćenje kretanja predavača i u nekim slučajevima, kod kompleksnijih sistema - prepoznavanje akcija korisnika.

Digitalizacija pisanih materijala je grupa svrstanih softversko-hardverskih komponenta za transformaciju rukopisa u elektronski format, odnosno direktno prikupljanje podataka preko pametnih tabli i ekrana osjetljivih na dodir (Xu et al, 2009). Dokazano je da se sistemi za automatizovani zapis materijala na predavanju u elektronski oblik pozitivno odražavaju na uspeh studenata (Sevindik, 2010). Digitalizovane beleške i zapisi se dalje elektronskim putem distribuiraju preko veb portala gde je moguć njihov pregled, dopuna, sinhronizacija sa audio i video sadržajima.

Sistemi za video/audio konferencije su komercijalno široko zastupljeni sistemi koji se sastoje od interaktivnih telekomunikacionih tehnologija za uspostavljanje istovremene dvosmerne komunikacije putem slike i zvuka. Primer sistema za video/audio konferenciju je tele-edukacija pomoću koje se pokušava pružiti iskustvo za predavače i slušaoce što sličnije tradicionalnom (Xie et al. 2001), koje predstavlja ključ uspeha za video predavanja na daljinu. Interakcija bez direktnog “fizičkog” kontakta nije adekvatna u primeni edukacije, ne samo zbog nedostatka povratne reakcije već zbog samog sociološkog faktora. Kao posledica nastaju sistemi koji imitiraju realnost (virtuelizacija učesnika), npr. virtuelne učionice i mešovito učenje na daljinu.

Virtuelizacija učesnika je stvaranje virtuelnog okruženja koje simulira prikaz tradicionalne učionice. Konkretno, virtuelizacija se može posmatrati kao softverski rendering virtuelnog prostora učionice, u koji se digitalno umeću slika ili video zapis studenta strimovan u realnom vremenu; ili su učesnici reprezentovani svojom virtuelnom zamenom - avатарom.

Navedene komponente u realnom okruženju se sistematski dopunjuju; a u određenim slučajevima i predstavljaju jedinstvenu celinu, zbog čega se teško u nekim sistemima razgraničava gde jedna kategorija počinje a druga se završava.

2.2. Komparivni pregled pametnih učionica

Stvaranje ambijenta tradicionalne učionice kod predavanja na daljinu, realizuje se kao digitalni prikaz studenata i predavača, dostupnih kroz jedinstven pogled za sve učesnike u predavanju, bez obzira gde se oni nalazili. Konkretno, ovaj dvostruki prikaz obuhvata: media-tablu (analognu tradicionalnoj tabli), na kojoj se bilo kakav zapis prikazuje na daljinskom klijentu koji koristi student; i studentovu tablu, koja se, ukoliko predavač odluči preslikava na svim ostalim klijentima (Xie et al. 2001). Prednost ovakvog sistema je što nije neophodno da predavač bude u blizini računara, dok svi učesnici mogu da se vide međusobno, prate predavanje i aktivno učestvuju u njemu. Svakom studentu je omogućeno da razgovara sa predavačem kao u normalnom okruženju. U radu (Xie et al. 2001) prikazano je iskustvo predavača nakon susreta sa pametnom učionicom koji nisu želeli da se vrate tradicionalnom načinu predavanja.

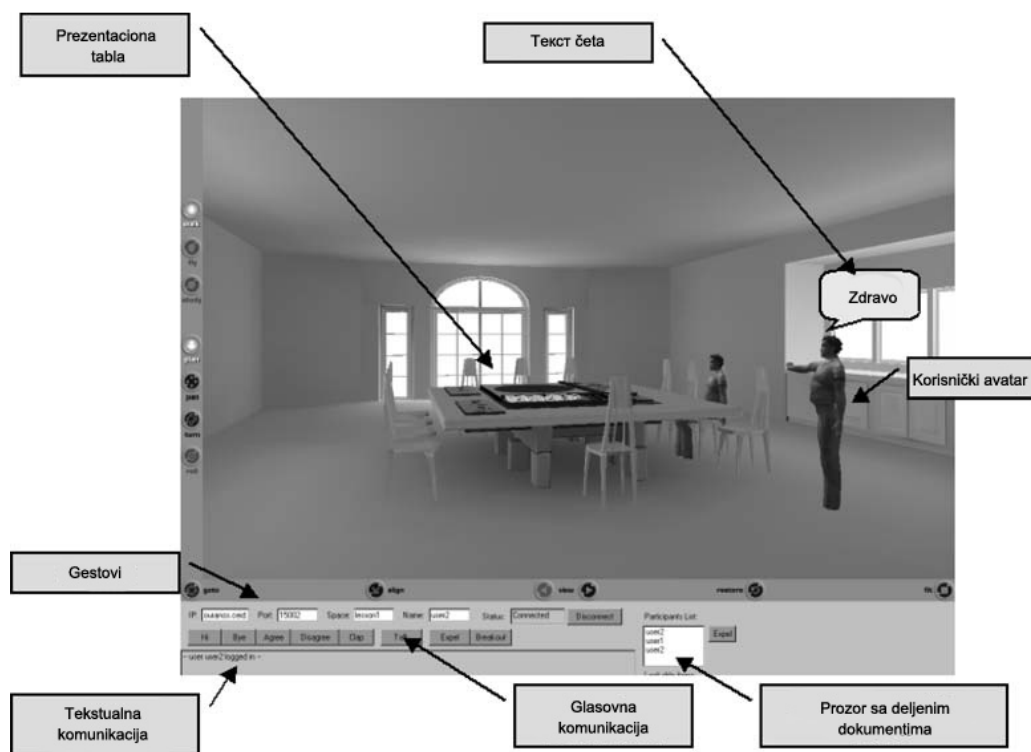
Standard Natural Classroom (SNC) je pametna učionica proširena tehnologijama za elektronsko učenje, s ciljem da se napravi interaktivno okruženje za lokalne i udaljene studente. Funkcionalnost ove učionice je dobar primer upotrebe različitog spektra tehnologija koje omogućavaju:

- detekciju i prepoznavanje lica pomoću adaboost (*Adaptive Boost*) algoritma i lokalnih binarnih paterna (LBP)
- praćenje različitih entiteta pomoću generalizovane metode pokreta (GMP) i osobina boja
- detektovanje pažnje studenata praćenjem pokreta oka
- prepoznavanje izraza lica pomoću LBP tehnike
- detektovanje predavačevih pokreta kao komandi pomoću *fuzzy* neuralnih mreža
- prepoznavanje glasovnih komandi pomoću HTK ili ViaVoice-a
- otkrivanje položaja osoba na osnovu niza mikrofona
- prepoznavanje emocija na osnovu kratkoročnih i dugoročnih osobina govora

Predavač u SNC može slobodno da se kreće za vreme prezentacije koristeći lasersku olovku, glasovne komande i pokrete. Studenti mogu da odluče da li će nastavu pratiti u učionici, SNC učionici, ili od kuće (Luo et al. 2009).

Virtuelizacija učesnika je postala idejno prihvatljiva usled ekspanzije igračke industrije i naučno fantastičnih filmova koji nude perspektivu treće dimenzije (3D). Može se pretpostaviti da će jednom predavanje na daljinu izgledati poput današnje posete bioskopu sa 3D projekcijom, gde će svaki od učesnika pratiti nastavu iz fizički udaljene lokacije.

Jedan primer kreiranja trodimenzionalnih virtuelnih prostora je platforma EVE (Bouras et al. 2006), koja koristi Virtual Reality Modeling Language (VRML) za prikaz 3D objekata, VRML External Authoring Interface (VRML-EAI) za implementiranje interfejsa između virtuelnih prostora i eksternih alata; i H.323 protokol za prenos zvuka. Predavanja u okviru EVE učionice se odvijaju u virtuelnom 3D okruženju koje se sastoji od stvarnih objekata koji se inače mogu naći u običnoj učionici (Slika 2).



Slika 2 Virtualizacija učesnika u predavanju

Problem dvosmerne komunikacije na daljinu je čest u ovoj oblasti. U radu (Xin et al. 2010) prikazani su komunikacioni protokoli, sistemski server, hardverski zahtevi i potreban protok u procesu implementacije sistema virtuelne učionice. VAVC sistem je baziran na klijent-server modelu, a okruženje se sastoji od klijenata i više servera koji komuniciraju preko Interneta. Okruženje i izgled scene u virtuelnoj VAVC učionici se modeluje u 3D Studio Max-u. Svi učesnici predavanja mogu da komuniciraju i dele znanja u virtuelnom ambijentu koji je karakterističan za video igre iz trećeg lica.

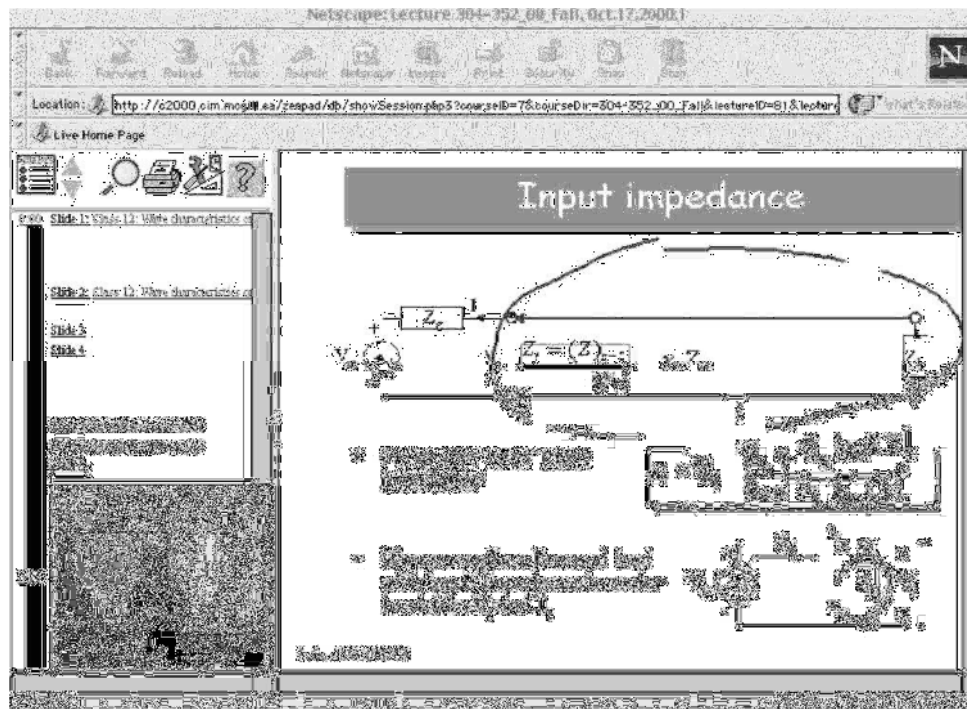
Istraživanje (Stiefelhagen et al. 2006) prikazuje modele za identifikovanje predavača, pomoću niza kamera i mikrofona. Prezentovani su algoritmi za prepoznavanje lica, orijentacije glave i fokusa pogleda. Kako bi se izbegla potreba držanja mikrofona, u neposrednoj blizini predavača smešten je niz mikrofona. Prepoznavanje reči je zahtevan proces zbog verovatnoće pojavljivanja akronima koji se ne koriste u književnom jeziku, nerazumljivih deonica, nedovršenog govora (kod interaktivnih predavanja), razlike u akcentima. Navedene poteškoće kod prepoznavanja reči stvaraju potrebu "treniranja" novih jezičkih i akustičnih modela. Prepoznavanje predavača je uvodna komponenta, ali se zavisno od funkcija koje slede, može rešiti na različite načine (npr. ukoliko logovanje nije zasnovano na glasovnom prepoznavanju; kontekst sistema je takav da se zna položaj predavača, i sl).

Pravac u kome je usmerena pažnja se može odrediti na osnovu usmerenosti pogleda, odnosno položaja glave. Eksperimentalni rezultati pokazuju da se veća tačnost kod određivanja položaja glave postiže korišćenjem video tehnike; ali kombinacijom većeg broja raspoloživih tehnika po pravilu ostvaruju se optimalniji rezultati (Tabela 1). Izuzetak je slučaj kada je predavač sniman audio i video tehnikama, izlaganje vodio u nedovoljno osvetljenom delu prostorije bez mnogo pokreta; što je rezultovalo velikim stepenom greške pri upotrebi video tehnike (Stiefelhagen et al. 2006). Prema tome, može se zaključiti da tačnost proporcionalno raste sa povećanjem kvalitativnih parametara uključenih u sam proces merenja, sa potrebom da se ispune specifični uslovi.

Tabela 2 Prosečna greška u (cm) pri proceni pozicije glave (Stiefelhagen et al. 2006)

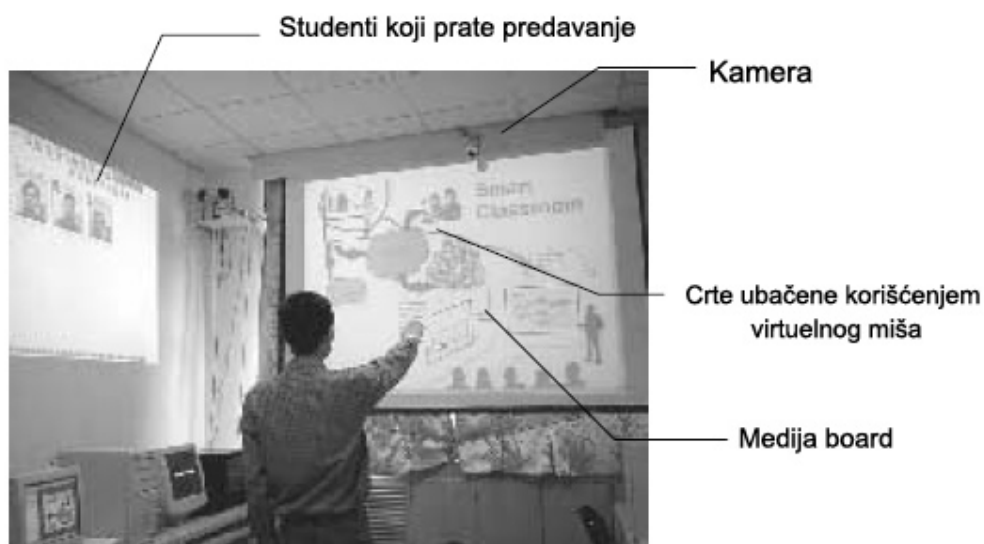
Način praćenja	Svi frejmovi	Govorne deonice
Audio	46.1	41.7
Video	36.3	36.5
Audio i video	30.5	29.1

Učionica opremljena mehanizmima za prepoznavanje pokreta i govora reaguje prema predavačevoj aktivnosti, bez potrebe za manuelnim upravljanjem. Kako bi se približnije objasnio ovaj scenario u daljem tekstu opisan je primer (Winer & Cooperstock, 2002). Kada se predavač uloguje na računar u učionici, sistem isključuje svetla, spušta ekran, pali projektor i prebacuje izlaz iz računara na projektor. Glavna ideja automatizacije ovih procesa je da se predavač usresredi na predavanje, umesto da vreme protекne u nameštanju opreme.



Slika 3 Multimedijalno predavanje

Pored automatizacije, učionica pruža mogućnost da se snime, spoje i sinhronizuju digitalne beleške i zapisi sa elektronske table (Powerpoint, beleške) sa audio/video zapisom predavača. Potom, sistem dozvoljava studentima da ocene predavanje, sa uvidom u prethodna predavanja preko veb interfejsa. Platforma omogućava studentima da urade analizu materijala sa predavanja. U rada je takođe istaknut značaj *feedback*-a u realnom vremenu, odnosno postavljanje pitanja od strane studenata u momentu prezentacije materije. Međutim, konstantna pitanja mogu ometati proces nastave i bolji se rezultati postižu ukoliko je auditorijum fokusiran samo na predavanje, a profesor prepozna koliko je predavanje razumljivo ili interesantno posredstvom tehnologije. Pasivno, studenti kada pregledaju materijal online mogu da koriste opciju prethodno postavljenih pitanja PAQ (Previously Asked Questions) za bilo koji slajd, ili da postavie novo pitanje za koje je moguće da naknadno dobiju odgovor. Nova pitanja se integrišu kasnije zajedno sa odgovorima kao nova PAQ (Winer & Cooperstock, 2002).



Slika 4 Tele-edukacija. Učenje na daljinu u realnom vremenu

Virtuele (laserske) olovke kao sastavni deo opreme jedne učionice koriste se za digitalizaciju pisanih materijala ili simulaciju miša - uPen (Xiaojun et al. 2006), Virtual pens (Xu et al, 2009). Laserski pointer opremljen tasterima i bežičnim komunikacionim

modulom, omogućuje predavaču potpunu slobodu kretanja pri upotrebi virtuelne olovke. Prapatni softver dozvoljava da se naprave profili za različite korisnike kako bi se omogućila upotreba većeg broja uređaja istovremeno.

Virtuelna režija (Xu et al, 2009) je naziv za grupu sistema koji se koriste za automatizovano snimanje predavanja, uz mogućnost praćenja kretanja predavača. Praćenje se obezbeđuje snimanjem kompletne scene preko širokougaone kamere koja navodi drugu kameru koja je bliža predavaču; ili snimanjem pomoću jedne kamere, tako što se predavač zadržava u centru scene na osnovu tehnika digitalnog kropovanja i mehaničkog praćenja (Yokoi & Fujiyoshi, 2005).

Presentations Automatically Organized from Lectures (PAOL) je rešenje koje snima sadržaj prezentacije prikupljanjem sekvenci slika visoke rezolucije (Dickson et al. 2008). PAOL snima seriju slika (frejmova), digitalno uklanja predavača sa scene, i sačuvava slike kada se pojavi stabilna verzija slike (koja je aktuelna duži vremenski period). Glavni proces se odvija u realnom vremenu brzinom od 1 frejma po sekundi (fps) po kameri; a sledeći proces koji obuhvata uklanjanje duplikata i povećavanja čitljivosti, započinje nakon snimanja predavanja, odvija se brzinom od 5 fps i završava najkasnije 2 minuta nakon završetka predavanja.

Open Smart Classroom (Suo et al. 2007) je pametna učionica opremljena veb servis tehnologijom kako bi se pružila bolja mogućnost nadogradnje i skalabilnosti, kao i da bi se izašlo u susret novim zahtevima učenja na daljinu. Ova vrsta pametne učionice bazirana je na Open Smart Platform-i, računarskoj infrastrukturi za inteligentna okruženja koja omogućava tri nove funkcionalnosti:

1. otvoreni standardni interfejs za povezivanje i komunikaciju sa mobilnim uređajima, bez naknadnih podešavanja
2. otvoreni servisni kanal za povezivanje internih modula i drugih sistema
3. otvoren pristup u kome više inteligentnih okruženja mogu da se povežu i komuniciraju međusobno.

Pojam *blending-reality* kod pametnih učionica predstavlja osnovu tele-edukacije nadograđenu virtuelnim prikazom učesnika. Blending-reality pametna učionica realizovana na univerzitetu u Kini (Di et al. 2008) obuhvata: interakciju čoveka i računara uz pomoć glasa čime se smanjuje potreba upotrebe tastature i miša; video/audio komunikaciju u realnom vremenu u oba smera, inteligentan video sistem sa praćenjem, kako bi predavač mogao slobodno da se kreće za vreme prezentacije; i snimak učionice, da bi učesnici mogli međusobno da se vide.

Praktična primena *blending reality* mogla bi izgledati na sledeći način: Podizanjem sistema inteligentni kontrolni modul na osnovu glasa vrši automatsko prepoznavanje korisnika. Kada predavač izgovori svoje ime, kontrolni modul učitava model govora, posle čega je moguće da se izdaju glasovne komande, npr. "otvori dokument", "strimuj prezentaciju", "sledeći slajd", itd. Predavač se slobodno može kretati dok izlaže, zahvaljujući glasovnim komandama i automatskim kamerama sa praćenjem. Zapis sa table se prenosi do studenata i prakazuje na jednom zidu učionice, a na drugom su prikazane slike studenata u realnom vremenu, sa ciljem da se približi ambijent tradicionalne učionice. Ukoliko student želi da zatraži reč - slika studenta će to signalizirati treperenjem.

Primena RFID se može pronaći u različitim oblicima a jedan primer je IBDM (Gu & Shen, 2010) bežična senzorska mreža koja se sastoji od dve vrste nodova. Prva vrsta ima RS-232 port za povezivanje sa računarom i postoji samo jedan ovakav nod u mreži, a druga vrsta se raspoređuje na više lokacija i komunikacija se odvija pomoću radio signala koji formiraju sićušnu bežičnu mrežu. Mreža se koristi za praćenje aktivnosti u prostoriji na osnovu emitovane toplote tela.

Rad (Ren & Hu) prezentuje rešenje za prepoznavanje akcija predavača kako bi se propratno doveli u funkciju virtuelni miš i virtuelni kameraman. Prepoznavanje pokreta se vrši preko hibridnog ljudskog modela (eng. Hybrid human model), na kome se primenjuje funkcija B-krive drugog reda kako bi se detektovala ramena i prikupile osnovne informacije o pokretima (uglovi laktova, parametri pokreta lica i obe ruke).

Potom se pomoću verzije skrivenog Markovljevog modela tzv. Primitive-based Coupled Hidden Markov Model (PCHMM) detektuju akcije zavisno od konteksta.

Tabela 3 Pregled pametnih učionica sa karakteristikama: [A] virtuelni asistenti; [R] automatizovano snimanje predavanja; [D] digitalizacija pisanih materijala; [K] sistemi za video/audio konferencije; [V] virtuelizacija učesnika

Naziv projekta	Funkcija	Osnovne karakteristike	Tehnike
Blending-reality smart Classroom (Di et al. 2008)	A, D, K, V	Interakcija čoveka i računara uz pomoć glasa; video/audio komunikacija u realnom vremenu u oba smera; inteligentan video sistem sa praćenjem; 3D prikaz scene učionice	Simon - softver za prepoznavanje govora; real time video capture VFW (<i>Video for Windows</i>) SDK; Vega 3D simulator
Eve platforma (Bouras et al. 2006)	V, D	Virtuelna predavanja sa audio konferencijom gde je svaki učesnik zamenjen svojim avатарom	Virtual Reality Modeling Language (VRML) za prikaz 3D objekata; VRML External Authoring Interface (VRML-EAI) za implementiranje interfejsa između virtuelnih prostora i eksternih alata; H.323 protokol za prenos zvuka
IBMD (Gu & Shen, 2010)	A	RFID bežična senzorska mreža	Detektovanje pokreta na osnovu promena temperature pomoću senzora
VAVC (Xin et al. 2010)	V, K	Virtuelna učionica, hardverski i tehnički zahtevi; modelovanje ponašanja avatara	3D Studio Max za modelovanje okruženja; <i>Virtools</i> DEV za interakciju avatara; 3DVIA <i>Player</i> za kretanje kroz scene; <i>Virtools</i> SDK (<i>Software Development Kit</i>) and <i>Virtools</i> VSL (<i>Virtools Scripting Language</i>) za komunikaciju sa serverom
PAOL (Dickson et al. 2008)	R	Automatizovano snimanje sadržaja table snimanjem serija slika	Dve kamere za snimanje prezentacije u rezoluciji 1024x410 px; slika se deli na 16x16px blokove a potom se konvertuju pikseli, remapiraju boje i digitalno uklanja predavač

Prepoznavanje pokreta (Ren & Hu)	A	Prepoznavanje akcija predavača kako bi se propratno doveli u funkciju virtuelni miš i virtuelni kameraman	<i>Primitive-based Coupled Hidden Markov</i> model (PCHMM) za detektovanje akcija zavisno od konteksta
uPen (Xiaojun et al. 2006)	A, D	Laserski pointer sa tasterima, mogućnost pisanja po površini i manipulacija na daljinu sa svim funkcionalnostima miša	Softver baziran na Smart Platformi koji obezbeđuje metode komunikacije i koordinacije, detekciju laserske tačke, emulaciju bežičnog miša, identifikaciju različitih olovki
Standard Natural Classroom (Luo et al. 2009)	A,D,K	Pametna učionica sa širokim spektrom komponenata za prepoznavanje lica, govora, pokreta, pažnje, emocija u glasu, izraza lica; i tehnologijama za tele-edukaciju, s ciljem da se napravi interaktivno okruženje za lokalne i udaljene studente	Detekcija i prepoznavanje lica pomoću <i>adaboost</i> algoritma i lokalnih binarnih paterna (LBP), praćenje različitih entiteta pomoću generalizovane metode pokreta (GMM) i osobina boja; detektovanje pažnje praćenjem pokreta oka; prepoznavanje izraza lica pomoću LBP, detektovanje pokreta kao komandi pomoću <i>fuzzy</i> neuralnih mreža; prepoznavanje glasovnih komandi pomoću HTK ili <i>ViaVoice-a</i> ; pronalaženje položaja osoba na osnovu niza mikrofona; prepoznavanje emocija pomoću kratkoročnih i dugoročnih osobina govora
Smart Remote Classroom (Shi et al, 2002)	K, R, D, V	Platforma za učenje na daljinu u realnom vremenu zasnovana na hibridnom protokolu aplikacionog sloja za višeznačno prekrivanje TORM i adaptivnoj šemi za isporuku sadržaja AMTM. <i>SameView</i> softver za pravljenje multimedijalnog sadržaja snimanjem predavanja sa embedovanom 3D tehnologijom.	TORM protokol dinamički kreira tunele radi povezivanja čvorova lociranih u različitim "ostrvima". Nakon primanja paketa kroz tunel, paketi se prosleđuju ostalim čvorovima u domenu. Šema ATMT pruža različite usluge za isporuku sadržaja, opisujući organizaciju multimedijalnog sadržaja kako bi se različiti formati podataka (slika, tekst, video) isporučili klijentima na osnovu potrebnog protoka i zahteva primaoca
	A, R, D	Pametna učionica opremljena	Interakcija predavača sa računarom i automatska

MC13 (Winera & Cooperstock, 2002)		mehanizmima za automatsku kontrolu ambijentom, snimanje i sinhronizaciju snimaka predavanja sa digitalnim zapisima prikupljenih preko elektronske table i tableta. Materijal je dostupan studentima preko vebe	kontrola postignuta na osnovnu funkcionalnosti nadograđene platforme <i>Classroom 2000</i> i pristupu implementacije reaktivne video-konferencijske sale
CyLab (Pishva, 2007)	K, R	Predavanje na daljinu između dve pametne učionice sa sistemom za automatsko snimanje prezentacije. Praćenje snimljenih predavanja na zahtev	Video-konferencija se prenosi na daljinu preko linka brzine 4 Mbps, protokol H.320/H.323. Rezervni link 512 Kbps ISDN. Prikaz u udaljenoj učionici preko <i>Microsoft NetMeeting-a</i> . Snimanje scene sa praćenjem predavača
Intelligent Environment for Tele-education (Xie et al. 2001)	A, K	Inteligentna okruženja koja prepoznaju govor i pokrete i vrše određene radnje na osnovu zapaženog. Tele-edukacioni sistem koji koristi <i>SameView</i>	Upravljanje pomoću glasa i pokreta, zamenjivanje funkcija miša. Akcije okruženja se izvršavaju na osnovu multimodalnog fraziranja. Svaka akcija se ispituje kao poseban modalitet dok se ne složi sa predefinisanim fraznom strukturom. Protokol TORM
CHIL (Stiefelhaven et al. 2006)	A	Rad prikazuje modele za detektovanje predavača, pomoću niza kamera i mikrofona i njegovo praćenje, prepoznavanje govora i orijentacije glave.	Signal iz većeg broja izvora (kamera i postavljenog niza mikrofona) se procesira različitim tehnikama radi utvrđivanja položaja predavača. Algoritam se zasniva na pretpostavci da je predavač fizički najaktivniji u prostoriji. Prepoznavanje govora na osnovu <i>Janus recognition toolkit</i> (JRTk) i lica koriste se na osnovu treniranih modela. Svi senzori i komponente rade pod NIST <i>Smartflow</i> platformom
Open Smart Classroom (Suo et al. 2007)	K	Pametna učionica opremljena veb servis tehnologijom za učenje na daljinu	<i>Open Smart Platform-a</i> , otvoreni standardni interfejs za povezivanje i komunikaciju sa mobilnim uređajima bez naknadnih

podešavanja; otvoreni servisni kanal za povezivanje internih modula i drugih sistema; otvoren pristup u kome više inteligentnih okruženja mogu da se povežu i komuniciraju međusobno

Jedno istraživanje (Sevindik, 2010) sprovedeno u pametnoj učionici pokazuje kakvu ulogu može da igra ovakvo okruženje u edukaciji. Pri realizaciji istraživanja, uzorak je nasumično izabran, a većina studenata su sa druge godine studija. Eksperimentalna grupa sačinjena od pedeset studenata predavanje je pratila u pametnoj učionici, a kao rezultat zaključeno je da postoji evidentan napredak u odnosu na studente koji su predavanja pratili u tradicionalnoj učionici. Pametna učionica u ovom istraživanju podrazumevala je okruženje opremljeno opremom za video konferencije, prenos predavanja uživo, automatizovano snimanje predavanja i zapisa na tabli. Testovi koje su studenti radili pre i posle predavanja odnosili su se na oblasti koje su bile sama tema predavanja. Eksperiment je započet sa osamdeset studenata, ali je ceo proces testiranja sproveden na svega šezdeset šest, od kojih je 20% bila kontrolna grupa. Ovaj uzorak je možda nedovoljno reprezentativan, ali ide u prilog upotrebi pametnih okruženja u edukaciji.

2.3. Platforme inteligentnih okruženja

Praktično je nemoguće sve navedene komponente i funkcionalnosti implementirati na jednom računaru zbog ograničenih resursa, što je i jedan od razloga pojave distribuiranih računarskih platformi za inteligentna okruženja. Ove platforme se mogu podeliti u dve različite grupe zavisno od strukturalnog modela:

1. distribuirane modalne platforme sa komponentama (npr. CORBA od OMG-a i DCOM od Majkrosofta)
2. multi-agentne platforme (OAA od SRI ili Aglets od IBM-a).

Kod prve grupe platformi postoji samo jedan izvršni proces u sistemu, a objekti se izvršavaju samo kada su pozvani od strane centralne logike. S druge strane multi-agentni model se uklapa u koncept sistemske strukture sa nezavisnim entitetima, razdvajanjem individualnih softverskih modula (agenata) koji imaju svoje izvršne procese, zbog čega su ove platforme praktičnije za upotrebu u inteligentnim učionicama.

2.4. Uloga društvenih nauka

Interaktivni procesi između predavača i auditorijuma u pametnoj učionici zahtevaju razumevanje sociološke i psihološke pozadine aktivnosti koje se prate. Analiziranje signala generalno zahteva analizu svih paterna koji okružuju vizuelnu i/ili verbalnu poruku i utiču na njeno značenje ili karakterišu neku sociološku ulogu.

2.4.1. Sociološki signali

Prva istraživanja na temu emocija pojavila su se 1962. godine sa teorijom Sylvan Tomkinsa, koja je potvrdila mnoge tradicionalne pretpostavke o izražavanju emocija i predložila jednu novu, koja je promenila način poimanja biheviorističke pozadine emocionalnih izraza. Prema Tomkinsu, svaka osnovna emocija varira u intezitetu i

sastoji se od jednog procesa u mozgu, čije pokretanje se odražava na skup komponenti: vokalne izraze, pokrete mišića lica i instrumentalne akcije (Russell et al. 2003). Emocije izražene u određenom socijalnom kontekstu, predstavljaju sociološke signale.

U edukaciji, sociološki signali studenata mogu da se koriste kao smernice koje bi preko povratnog kanala u realnom vremenu ukazivale na nedostatke predavanja, stepen razumevanja prezentovane materije i zainteresovanost auditorjuma. Oblast koja se bavi istraživanjem ljudskih socioloških signala s ciljem da se tehnologija unapredi u prirodnijem i fleksibilnijem pravcu naziva se Social Aware Computing. U većini slučajeva verbalna komunikacija je praćena ovim signalima, koji se nazivaju još i iskreni signali zbog svog podsvesnog karaktera, čime se odaju informacije kao psihološka pozadina određenih postupaka.

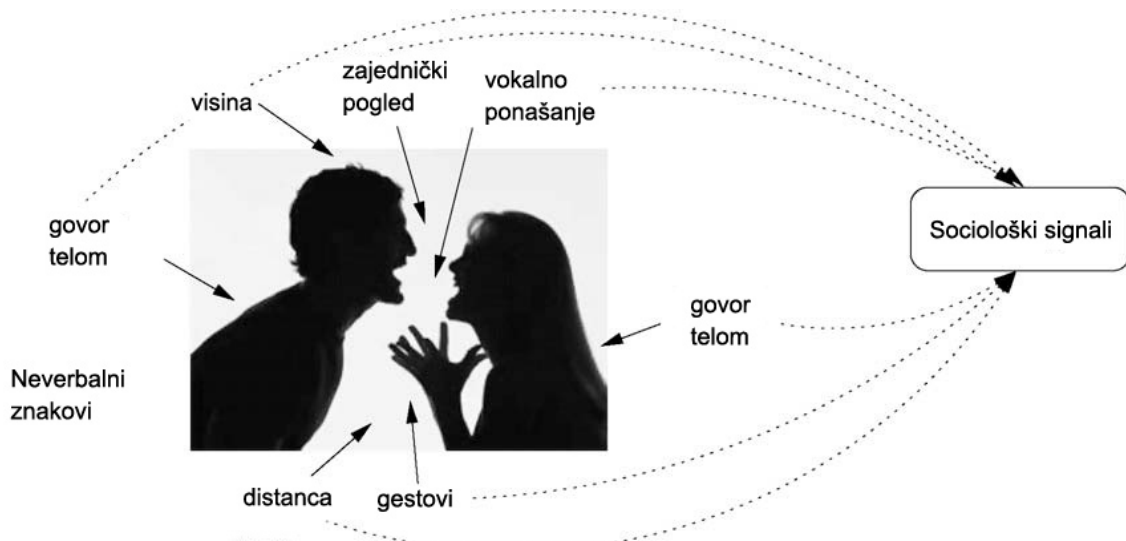
Sociološki signali su frekventna pojavljivanja socioloških znakova, i definišu se kao privremene promene neuro-mišićne i neuro-psihološke aktivnosti (Ambady & Rosenthal, 1992) (Pesarin et al. 2010). Različita ljudska ponašanja se mogu pouzdano predvideti praćenjem pokreta lica, mimike, aktivnosti, isticanja i sinhronizacije (Pentland, 2008) između navedenih. Ovi urođeni signali potiču još iz ranijeg perioda evolucije, kada su se pojavili zajedno sa prvim vidovima komunikacije.

Pri interakciji neverbalna komunikacija je izvor pouzdanih signala koji odaju informacije o emocijama, mentalnim stanjima, ličnosti, stavovima, i drugim osobinama. Gotovo je nemoguće prikriti ove signale, jer su podsvesni, a pokušaj njihovog imitiranja upadljiv je za posmatrača. Primenom senzora i adekvatnih algoritama, neke od signala je moguće zabeležiti sa visokim procentom tačnosti.

2.4.2. Analiza socioloških signala

Nekoliko istraživanja pokazalo je vezu između vokalnih osobina glasa, mimike lica i pokreta ruku, na osnovu čega je hipotetički zaključeno da se jedan impuls istovremeno može podsvesno odražavati preko većeg broja modaliteta: promenama lica, glasa i

pokreta tela (Pentland, 2006). Procesiranje socioloških signala ima za cilj razvoj teorija i algoritama koji mogu da prikažu kako se ljudska bića ponašaju za vreme interakcije, spajajući znanja iz oblasti sociologije, psihologije i računarskih nauka (Pantic et al. 2009) (Pesarin et al. 2010).



Slika 5 Sociološki signali

Elementi koji imaju udela u sociološkom kontekstu se mogu podeliti na sledeće kategorije (Vinciarelli et al. 2009):

1. fizički izgled
2. gestikulacija i držanje
3. Izrazi lica i pokreti očiju
4. vokalni signali
5. okruženje i prostor

Fizički izgled

U grupu signala koji karakterišu fizički izgled spadaju: visina, oblik tela, boja kože i kose; i veštačke karakteristike: odeća, šminka, ornamenti i sve što utiče na izgled lica i tela. Fizički izgled nije sam po sebi sociološki signal, ali može da oda značajne

informacije o ličnosti osobe koja se prati. Pre svega ovo se odnosi na karakter jedne osobe, i način na koji kod svake mozak prima i obrađuje signale, na osnovu čega razlikujemo dve vrste osoba: introvertne i ekstrovertne; o čemu će biti više reči kasnije u radu.

Gestikulacija i držanje

Prema pregledu istraživanja iz oblasti (Vinciarelli et al. 2009), pokreti tela menjaju se u skladu sa emocijama. Osnovna istraživanja su pokazala da gestovi poput pognute glave, dodirivanje lica i promena u držanju tela mogu da odražavaju neprijatnost ili sramotu. U većini slučajeva gestovi su propratni deo interakcije koji naglašava određenu radnju. Međutim, u određenim slučajevima gestovi su podsvesni i predstavljaju “iskrenu” informaciju.

Signali poput adaptatora otkrivaju dosadu, stres i negativna osećanja prema drugima, manifestuju se najčešće podsvesno (npr. češanje, dodirivanje nosa i uha, uvrtnje kose), manipulacija malim objektima (npr. igranje sa olovkom i papirom), ili samo-zaštitnički gestovi (npr. prekrštene ruke ili ritmično pomeranje nogu). Za pokrete se najčešće može reći da su najpouzdaniji patern za određivanje stava individue o sociološkoj situaciji u kojoj se nalazi.

Izrazi lica i pokreti očiju

Cilj sagledavanja paterna koji se izražavaju preko lica je identifikovanje određenog pokreta sa značenjem koje on reflektuje. Kod praćenja pokreta lica postoje dva pristupa: deskriptovanje na osnovu poruke i na osnovu znaka. Deskriptori koji se najviše upotrebljavaju pri proceni pokreta lica su strah, tuga, sreća, gađenje, i iznenađenje. Prema Ekmanu emocije se univerzalno prepoznaju preko pokreta mišića lica.

Tehnika koje se dosta primenjuje za manuelno mapiranje pokreta lica je Facial Action Coding System (FACS). FACS povezuje promene izraza lica sa aktivnostima mišića koji proizvode te promene. Definiše 9 različitih akcijskih jedinica (acion unit – AU) gornjeg

lica, 18 donjeg dela lica, 11 pozicija glave, 9 pozicija očiju i 14 dodatnih deskriptora za različite akcije. AU su najmanji vidljivi pokreti lica koje je moguće pratiti. Korišćenjem FACS moguće je kodirati praktično bilo koji pokret, dekompozicijom na jedinice koje proizvode pokrete. Ova tehnika se može upotrebiti za prepoznavanje socioloških signala: amblema (interaktivni signali: namigivanje), regulatora (konverzacioni mediator: klimanje glavom, razmena pogleda) i ilustratora (signala koje prate govor: podignute obrve).

Vokalni signali

Vokalno neverbalno ponašanje se odnosi na sve izgovorene tragove koji okružuju verbalnu poruku i utiču na njeno značenje. Efekat vokalnog neverbalnog ponašanja je evidentan kada je ton poruke ironičan. U ovom slučaju, licem se nagoveštava suprotno značenje od izgovorenog; i obrnuto.

Vokalno neverbalno ponašanje uključuje 5 komponenti: kvalitet glasa, jezičku i nejezičku vokalizaciju, tišinu, i paterne smene slušaoca i govornika. Kvalitet glasa se odlikuje visinom tona, tempom i energijom. Percepcija kvaliteta glasa odnosi se na način na koji je nešto izgovoreno (Crystal, 1969). Emocije poput besa ili straha često su praćene energetsom eksplozijom (Scherer, 2003). Tečnost govora karakterišu visok ritam i nedostatak pauze (Scherer, 1979).

Jezička vokalizacija uključuje sve zvukove koje nisu reči, kao da su zapravo reči (“hm”, “aha”, itd). Na ovaj način se zamenjuju reči koje se iz nekog razloga ne mogu pronaći, na primer kada neko na pitanje odgovori sa “hm”. Jezička vokalizacija se prepoznaje prema nekozistentnošću i najčešće označava neku vrstu otežavajuće okolnosti pri interakciji (Glass et al. 2008).

Drugi slučaj pojave ovih “reči” je u vidu tzv. povratnog kanala, gde se koriste da bi se upotpunio nečiji govor. U ovom slučaju se radi o izražavanju pažnje, odobravanja,

čuđenja, kontradiktornosti, kao i pokušaja da se preuzme vodeća reč (Shrout & Fiske, 1981).

Nejezička vokalizacija uključuje neverbalne zvukove: smejanje, jecanje, plakanje, šaputanje, uzdisanje, koji mogu a ne moraju da budu praćene rečima, tako da pružaju određene informacije o stavu pojedinca o određenoj sociološkoj interakciji. Na primer smeh nagrađuje poželjno sociološko ponašanje (Keltner & Haidt, 1999).

Istraživanja u psihologiji pokazala su da je slušalac u mogućnosti da razume osnovne emocije kao i neke složenije sociološke signale poput zabrinutosti, nervoze, dosade i seksualne zainteresovanost - pomoću smeha, kašlja, zevanja i drugih znakova (Russell & Fernandez-Dols, 1997).

Tišina se tretira kao jednostavan nedostatak govora, ali ustvari igra ključnu ulogu u vokalnoj analizi (Zellner, 1994). Kod govora javljaju se tri vrste tišine: tišina oklevanja, psiholingvistička tišina i interaktivna tišina (Richmond & McCroskey, 1995). Tišina oklevanja javlja se kada sagovornik ima problema pri izlaganju, npr. ako je tema zahtevna ili pri suočavanju sa suprotnim stavom u publici.

Psiholingvistička tišina javlja se kada govornik razmišlja o svojim sledećim rečima, zbog čega se ova vrsta tišine može tumačiti kao poteškoća ili problem upuštanja u konverzaciju. Interaktivna tišina može nositi informaciju o mestu na kom se odvija interakcija: tišina može da bude znak poštovanja prema sagovorniku, ignorisanja osobe kojoj se ne želi odgovoriti, kao i način da se privuče pažnja drugih formi komunikacije (pogleda ili izraza lica).

Još jedan bitan aspekt vokalnog neverbalnog ponašanja je preuzimanje glavne reči pri komunikaciji (Psathas, 1995). Ovo uključuje dve glavne komponente: regulaciju konverzacije i koordinaciju ili nedostatak nje za vreme menjanja govornika. Regulaciju u razgovoru uključuje ponašanja koja teže da se održi, prepusti, negira, ili zatraži glavna reč. Usmerenost pogleda i kvalitet glasa (kašljanje) koriste se za prepoznavanje

promene (Yule, 1996). Kada se radi o vokalnim neverbalnim paternima kao regulatorima konverzacije, određeni paterni visine tona i energije pokazuju nameru da se prepusti glavna reč. Takođe, jezička vokalizacija često se koristi kao forma da se na indirektan način zatraži reč.

Drugi bitan aspekt kod preuzimanja inicijative pri konverzaciji je koordinacija smene govornika (Burgoon et al. 1995). Interakcija u kojoj je čekanje između smene govornika predugačko ima tendenciju da bude definisana kao “neuspela“ konverzacija. Razlog je to što pri tečnoj konverzaciji uzajamna pažnja učesnika rezultuje sinhronizacijom promene govornika, tako što njihovi učesnici efikasno prepoznaju momenat kada je došao red da se prepusti ili preuzme reč. Preklapanja deonica govora u konverzaciji je demonstracija dominacije ili svađe (Smith-Lovin & Brody, 1989). Međutim, procenat preklapanja govornih deonica je uobičajena pojava i kreće se do 10% u svakodnevnom razgovoru (Shriberg et al. 2001).

2.4.3. Ekstrakcija socioloških signala

Ton glasa (Lepri et al. 2010) je specifičan za svakog pojedinca, ali njegova postojanost, odnosno varijacije, pružaju dodatne informacije o usresređenosti i podložnosti na uticaj od strane sagovornika. Ton glasa se analizira merenjem formantne (rezonantne) frekvencija glasa, promenljivosti formantne frekvencije, vrednosti autokorelacije vrhova, broja velikih autokorelacionih vrhova i njihova lokacija, promene energije u vremenu, itd.

Konverzaciona aktivnost se definiše energijom u frejmu, dužinom glasovnih segmenata, dužinom govornih segmenata i govornog inteziteta. Ona je obično pokazatelj zainteresovanosti ili angažovanja.

Uticaj predstavlja iznos uticaja koji jedna osoba ima na drugu pri interakciji, a izračunava se kao odnos govornih deonica koje se preklapaju u celini. Uticaj se u ovom

slučaju može tumačiti kao signal za dominantnost, a snaga uticaja za vreme konverzacije kao indikator pažnje.

Mimika za vreme interakcije se javlja kao oponašanje gestikulacije ili pokreta, praćeni kratkim izrazima: “da”, “aha”; ili razmena kratkih fraza: “razumeš?”, “razumem”. Mimika je kompleksno ponašanje zbog čega je veoma teško identifikovati paterne iz okruženja. Detektovanje mimike se do određene mere može odrediti preko Z-score modela upotrebljenog na kratkim promenama izgovora (ispod jedne sekunde).

Gestikulacija (pokreti tela, glave i ruku) se može identifikovati MHI (Motion History Images) tehnikom (Chippendale, 2006), koja koristi prepoznavanje kože i pokreta u vremenu radi ekstrakcije pokreta koji se ponavljaju na slikama i dodeljivanja energetske vrednosti takvim pokretima. Naglašeni pokreti imaju veće vrednosti. Gestikulacija u ovom slučaju se ekstrahuje za svaki frejm na frekvenciji od tri herca.

Problem tehničke analize socioloških signala obuhvata procese: snimanja scene korišćenjem različitih senzora, detekciju osoba u sceni pomoću različitih algoritama, ekstrakciju prikazane aktivnosti i njenu interpretaciju i klasifikaciju prema ciljnim kategorijama.

2.4.3.1. Prikupljanje podataka

Prikupljanje podataka se odnosi na korišćenje različitih vrsta senzora kako bi se snimile sociološke interakcije koje se odvijaju u realnom životu.

Izbor IoT uređaja i način i mesto postavljanja u određenoj postavci može da ograniči spektar paterne koje je moguće prikupiti. Na primer, nije moguće prikupiti informacije o pogledu, ako određen detektor nije uključen u postavku sistema. U senzore koji se

najviše koriste spadaju mikrofoni i kamere, i njihov broj i postavka zavisi od kompleksnosti ponašanja koje je potrebno zabeležiti i kasnije prepoznati kao i od zahtevanog kvaliteta digitalnog zapisa.

Prikupljanje podataka pomoću IoT uređaja tako da subjekat nije svestan da je posmatran, je metoda koja donosi mnogo tačnije rezultate, ali je ovaj proces daleko zahtevniji. Ovo se pre svega odnosi na zahtev postavljanja IoT uređaja, koji je potrebno da budu implementirani u okruženju a ne postavljeni na studente. Ukoliko se senzori postave na studente, uticaće se na prirodnost interakcije i na samu sposobnost studenta da se fokusira na predavanje.

2.4.3.2. Fuzija podataka iz više senzorskih izvora

U proteklih nekoliko godina istraživanja na polju prikupljanja podataka iz više senzorskih izvora je dobila na značaju kako za vojne primene tako i uopšte. Fuzija podataka predstavlja kombinovanje podataka iz više izvora i relevantnih informacija iz baza podataka kako bi se povećala tačnost u odnosu na upotrebu informacija koje dolaze od jednog senzora. Ovaj koncept nije nov već se bazira na prirodnom procesu prikupljanja informacija koji se kod ljudi i životinja manifestuje kao upotreba većeg broja čula da bi se uobličile informacije o nekoj pojavi. Kao najjednostavniji primer može se navesti proces određivanja kvaliteta hrane, kada je potrebno upotrebiti više čula – čulo mirisam, ukusa, itd, da bi se utvrdilo da li je hrana pokvarena ili ne (Hall & Llinas, 1997). Prema tome, fuzija informacije iz više izvora se prirodno odvija kod ogromnog broja procesa pri interakciji čoveka i životinja sa okruženjem da bi se identifikovale pretnje i povećave šanse za opstankom. Na sličan način potrebno je prikupiti informacije koje potiču iz više izvora i uraditi njihovo procesiranje kako bi se odredila pažnja studenata, odnosno njihova aktivnost tokom nastave upotrebom tehnologije Interneta inteligentnih uređaja.

2.4.3.3. Detekcija ljudi sa digitalnog signala

Senzori pri prikupljanju podataka se mogu upotrebiti tako da izlazni signal odmah daje informacije koje određuju sociološki signal ili određeno ponašanje. U nekim slučajevima, signal koji pripada jednom subjektu selektivno se razgraničava na mestu njenog nastanka. Međutim, ovo je slučaj kada se senzori, npr. elektrode postavljaju na subjekta. Neinvezivna postavka sistema dobra je praksa zbog kvaliteta informacija koje se dobijaju na ovaj način ali samo u sociološkom smislu.

Signal prikupljen na ovaj način gde se više od jednog pojedinca snima pomoću senzora, sadrži višak informacija (pozadinski šum), ili preklapanja različitih deonica.

Proces detekcije osobe kada se više subjekata nalazi na jednom zapisu naziva se segmentacija govornika i sastoji se od deljenja govornih zapisa na intervale koji pripadaju jednom glasu, automatskim prepoznavanjem (Shriberg et al. 2001). Segmentacija govornika se deli na sledeće faze: segmentacija podataka na segmente bez govora, detekcija smene govornika i tzv klasteringa (eng. Clustering) odnosno grupisanje govornih deonica prema jednom subjektu. U nekim slučajevima, ne očekuje se tišina pri konverzaciji tako da prvi korak nije neophodan.

Sistemi koji ne obuhvataju segmentaciju deonica sa i bez govora nazivaju se sistemi segmentacije govornika. Segmentacija deonica bez i sa govorom se najčešće izvršava korišćenjem mašinskih algoritama za učenje, treniranih pomoću različitih audio klasa (deonica bez govora na primer može da sadrži muziku, pozadinsku buku, tišinu, itd.). Tehnike koje se najviše upotrebljavaju u ovu svrhu su: veštačke neuralne mreže (artificial neural networks) (Ajmera et al. 2003), k nearest neighbours (Lu et al. 2002), Gaussian mixture model (Gauvain et al. 1998), itd.

Karakteristike koje se mogu ekstrahovati iz bilo kog signala su energija i autokorelacija (Rabiner & Schafer, 1978); a karakteristike koje se obično koriste kod prepoznavanja

govora: Mel frequency cepstrum coefficients (MFCC), Linear predictive coding (LPC), itd (Huang et al. 2001).

Prepoznavanje smene govornika se vrši deljenjem govornih segmenata na kratke intervale (npr. 2-3 sekunde) i merenjem razlika između dva uzastopna intervala: najveća razlika odgovara smeni govornika. Ovaj pristup je zasnovan na činjenici da zapis obuhvata bar dva subjekta. Ako ovo nije slučaj, jednostavna intonacija ili pozadinska buka može biti registrovana kao smena govornika.

Promena se može proceniti modelovanjem svakog intervala korišćenjem jedne Gausove metod, a razlika se račina simetričnom Kullback-Leibler divergencijom (Barras, 2004). Poslednji korak je segmentacija ili klastering – grupisanje segmenata koji pripadaju jednom glasu u jedinstven klaster. Ovaj proces se najčešće obavlja kroz iterativne pristupe (Barras, 2004) gde se klasteri inicijalizuju korišćenjem intervala između promene govornika detektovana u prethodnom koraku (svaki interval se konvetuje u set vektora preko tehnika za procesiranje govora (Huang et al. 2001) (Rabiner & Schafer, 1978), a potom se interativno spajaju na osnovu sličnosti modela koji ih reprezentuju (Gausova metoda ili GMMs). Proces spajanja se prekida kada se zadovolji uslov (broj sličnosti počinje da opada).

Pristup koji se najviše koristi za integraciju ovih koraka kao jedinstvenog frejmvorka je skriveni Markovljev model ili dinamična Bajesova mreža koja poravnava osobine vektora ekstraktovane u određenom vremenskom periodu (npr. 30 ms) i sekvenci stanja koje odgovaraju govorniku (Vinciarelli etl al. 2009).

2.4.3.4. Detekcija socioloških signala iz gestova

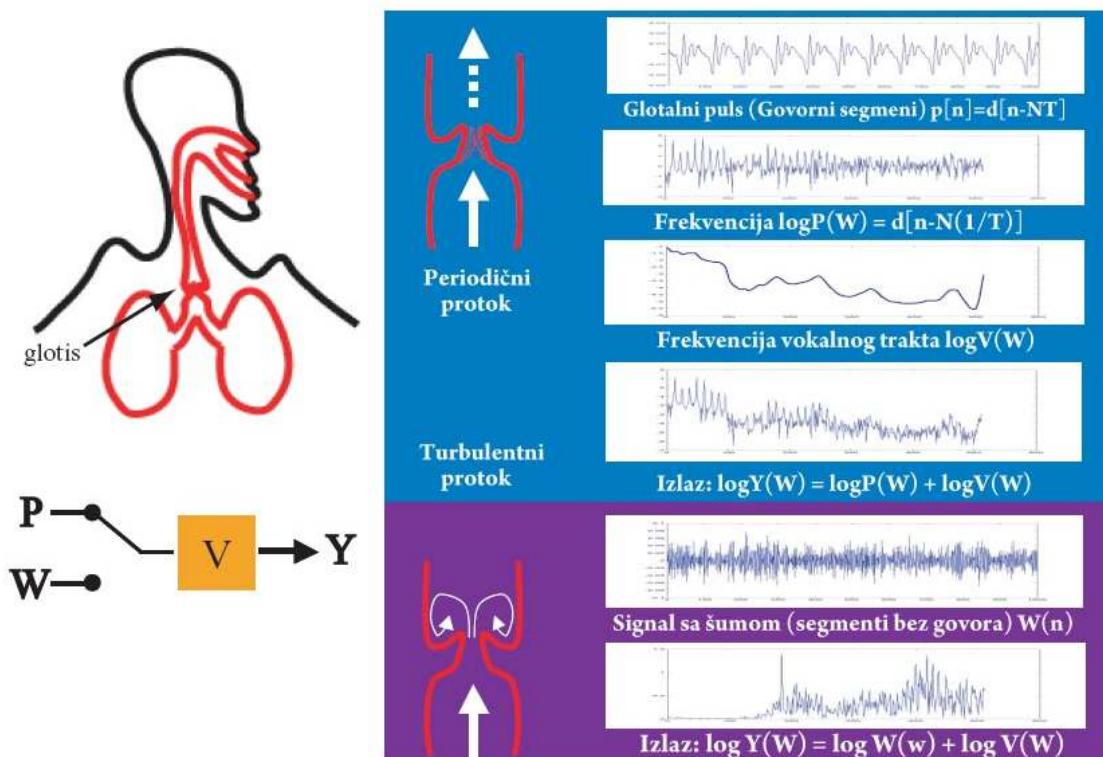
Istraživanje gestova se uglavnom orijentiše na njihovu primenu u procesu upravljanja, kao alternativa mišu i tastaturi, ili za automatsko čitanje jezika znakova. Postoje dva glavna izazova kod prepoznavanja gestova: identifikacija delova tela koji su uključeni u stvaranje gesta, a uglavnom se ruke koriste pri gestikulaciji; i modelovanje dinamike pokreta u vremenu.

Problematika identifikacija delova tela u dosadašnjim istraživanjima rešena je putem tehnika koje se odnose na izbor odgovarajućih vizuelnih karakteristika, npr. histogram orijentisanih gradijenta (histograms of oriented gradients – HOG), optički tok, spatio-temporal salient points i space-time volumes. Modelovanje dinamike pokreta u vremenu se primenjuje putem dinamičkog vremenskog iskrivljenja (Dynamic Time Warping - DTW) (Oikonomopoulos et al. 2006), skrivenog Markovljevog modela, itd (Sminchisescu et al. 2006).

Prepoznavanje hoda se može vršiti različitim metodama, od kojih neke zagovaraju prepoznavanje na osnovu zvučnih paterna. Međutim na ovaj način se ne može ekstrahovati svako sociološko ponašanje. Osnovni pristup kod prepoznavanja hoda je segmentacija siluete ljudskog tela na individualne komponente (noge, ruke, torzo, itd), a nakon toga se prikazuje njihova geometrija za vreme aktivnosti hodanja preko vektora distance, simetričnih operatora, geometrijske karakteristike tela i koraka (npr. razlika između glave i stopala ili karlice) (Bobick & Johnson, 2001), itd.

2.4.3.5. Detekcija socioloških signala iz vokalnog ponašanja

Pre nego se pristupi analizi socioloških signala iz vokalnog ponašanja, radi transparentnijeg pregleda trenutnog istraživanja na ovom polju, biće objašnjen proces izgovaranja glasova kod čoveka na uprošćenom modelu sa slike 6.



Slika 6 Govorni aparat čoveka za proizvodnju glasova. Pluća guraju vazduh kroz glotis da bi stvorila periodični puls, pri čemu se glotis otvara ili zatvara, ili ostaje otvoren kod turbulentnog toka vazduha. Rezultujući spektar se dalje oblikuje vokalnim traktom $V(W)$

Govor se može podeliti na dva tipa: glasovni i neglasovni. Glasovni govor je onaj koji poseduje ton, odnosno onaj za koji se može reći da je najbliži vokalima po karakteristikama. Neglasovni govor je sve ostalo (glasovi poput usnenih, npr “p”, zatim “s”, “š”, itd.). Ovo se delimično bazirano na činjenici da je broj vokala najzastupljeniji u regularnom govoru, jer je jezička struktura takva da oni predstavljaju većinu glasova u

svakoj reči. Za vreme glasovnih govornih segmenata, pluća stvaraju vazdušni pritisak u glotisu, koji u se otvara u određenom momentu i stvara put vazduhu, nakon čega se ponovo zatvori. Ovaj proces se odvija u ograničenom vremenskom periodu i amplituda ovakvog signala gotovo je impulsnog oblika $p[n]$. Furijerovom transformacijom nad ovih signalom $P[w]$ takođe se dobija signal impulsnog oblika sa periodom koji predstavlja ton signala. Signal potom putuje kroz vokalni trakt koji filtrira zvuk $V[w]$ u domenu frekvencije, što rezultuje kombinovanim izlazom $Y[w]$, koji predstavlja vokalni trakt pomnožen sa signalom (Basu, 2002):

$$Y [w] = V [w] * P[w]$$

Govorni aparat čoveka poseduje visok nivo fleksibilnosti na koji može da se manipuliše vokalnim traktom da bi se proizveo širok spektar rezonanci koji se nazivaju formanti.

U slučaju neglasovnog govora, pluća stvaraju pritiska dovoljno samo da bi se glotis otvorio i bio otvoren kao što je prikazano na desnom delu slike 6. Zvuk se dalje oblikuje zavisno od konfiguracije vokalnog trakta, gde spadaju i pozicija jezika i zuba. Na ovaj način nastaju glasovi “s” i “š”. Glasovi poput “p” i “t”, nastaju od pritiska koji se stvara i nestaje na drugim mestima vokalnog trakta (usana za glas “p” i jezika i zuba za glas “t”). Zajednička karakteristika za sve navedene slučajeve je da rezultujući zvuk nije periodičan.

Kompleksnije verzije vokalnog ponašanja se pored gestova sa sociološkog aspekta teško identifikuju. Različiti izrazi koji se smenjuju pri interakciji zahtevaju pre svega klasifikaciju prema izvoru, kako bi se prvo identifikovali učesnici interakcije, a zatim praćenje i prepoznavanje karakterističnih paterna vokalnog ponašanja.

Smeš kao jedna vrsta vokalnog ponašanja detektuje se primenom binarne klasifikacije, kao što je npr. metoda potpornih vektora (SVM) i tehnike koje se često koriste kod prepoznavanja govora Mel frequency cepstral coefficients (MFCC); ili perceptualno linearno predviđanje sa Gausovom metodom i neuralnim mrežama. Pomenuta istraživanja se odnose samo na audio signale; međutim, nekoliko mali je broj istraživanja koja su pokušala da realizuju prećenje audio-vizuelno prepoznavanje

nelingvističkih vokalnih signala. Detektor smeha kombinuje izlaz audio detektora koji koristi MFCC karakteristike govora i vizualnog detektora koji prepoznaje prostorne lokacije karakterističnih tačaka lica (Ito et al. 2005).

2.4.3.6. Dinamika modelovanja signala u vremenu

Modelovanje signala u vremenu je ključno za razumevanja bilo kakvog sociološkog konteksta. Bihevioristički paterni u vremenu, njihova korelacija i brzina izvođenja (npr. veze između pokreta mišića lica koji formiraju izraze lica), predstavljaju izazov, zbog čega postoji samo nekoliko studija o mogućnosti razlikovanja spontanih i namernih gestikulacija (pokreti obrva i osmeha baziranih na vremenskoj dinamici ciljnog pokreta lica, glave i ramena; multimodalna analiza audio i vizuelne dinamike ponašanja radi prepoznavanja emocija. Kao što je navedeno u (Vinciarelli et al. 2009) postojeće metode ne mogu da se nose sa dužim vremenskim skalama i analizom zavisnom od konteksta u robustnijim sistemima.

2.4.4. Metode segmentacije i klasifikacije socioloških signala

Proces određivanja uspešnosti predavanja je težak i za veoma iskusne predavace. Klasifikacionim šemama i segmentacijom moguće je određene sociološke signale kategorizovati u različite klase. Na primer, digitalni zapis zvuka se može sastojati od nekoliko klasa – govora praćen pozadinskom bukom ili šumom. Prema tome, metoda segmentacije audio zapisa se može upotrebiti da bi se utvrdilo gde govor i pozadinski šum počinju. Segmentacijom se takođe poboljšava metoda klasifikacije iz razloga tako što se prvo utvrde deonice prelaza između različitih klasa, gde je klasifikacija podložna greškama (Andersson, 2004).

Zhang (Lu et al. 2002) je opisao šemu za klasifikaciju koja se sastoji iz dva koraka:

- (1) diskriminacija između govornih deonica i deonica bez govora;

(2) glasovni segmenti se dalje procesiraju traženjem promena govornika, i deonice bez govora se dalje razvrstavaju od zvukova u pozadini i tišine. Tačnost ovog pristupa se kreće oko 96%,

Međutim, povećavanjem broja klasa smanjuje se i tačnost klasifikacije, zbog čega je preporučljivo maksimum sedam klasa (Li et al. 2001) . Tačnost segmentacije je visoko zavisna od kvaliteta audio signala tako da se procenat greške kreće između 14-1%.

Jedan od poznatijih klasifikatora sa nekomercijalnom licencom u javi je Weka (Hall et al. 2009). Weka je Java alat za mašinsko učenje koji utvrđuje sličnost između instanci, na primer ako se mere vrednosti i smeštaju u bazu redom, moguće je istrenirati algoritam koji će sa određenom tačnošću da pokazuje stepen pažnje u učionici.

Biblioteke za prepoznavanje govora Sphinx (Sphinx, 2013) i analizu zvučnog spektra JTransforms (Sphinx, 2013) imaju zadovoljavajuće individualne karakteristike, ali veliki broj pojedinačnih alata može dodatno zakomplikovati implementaciju.

2.4.5. Psihologija

U psihologiji Trejtova teorija je glavni pristup pri proučavanju ličnosti jedne osobe. Osobine se prema Trejtu mogu definisati pomoću ponašanja, misli i emocija (Huang et al. 2001). Osobine ekstroverzije i introverzije su centralna dimenzija ličnosti jedne osobe. Ekstrovertne ličnosti su dinamičnije, druželjubivije, dok su introvertne više rezervisane i manje su sklone socijalnim aktivnostima. Po pravilu introvertne ličnosti ne moraju da budu i usamljeni pojedinci.

Termin introverzije i ekstroverzije prvi put pominje Karl Jung (Barras, 2004). Bitno je napomenuti da svi sveobuhvatni modeli ličnosti u psihologiji i sociologiji koriste ovaj koncept, što ga čini univerzalnim za primenu u ovom istraživanju. Prema Jungu, introverzija i ekstroverzija se odnose na pravac fizičke energije. Ako se fizička eneregija pojedinca “kreće” van njega, onda je ta osoba ekstrovertna, a ako se kreće u suprotnom smeru onda je introvertna (Oikonomopoulos et al. 2006). Ekstrovertne

osobe osećaju povećanje primljene energije pri interakciji sa velikom grupom ljudi, ali i smanjenje energije kada ostanu sami.

Kod ove dve vrste osoba postoje razlike u ponašanju. Prema istraživanju (Sminchisescu et al. 2006), ekstrovertne osobe nose dekorativniju garderobu; introvertne, sa druge strane nose praktičnu, udobnu garderobu. Jedno drugo istraživanje (Sminchisescu et al. 2006), pokazalo je da ekstrovertne osobe preferiraju energičniju muziku od introvertnih.

Mnoge individue poseduju osobine koje ih svrstavaju u jednu od dve prethodno navedene kategorije; ali i svaka osoba može da ima odlike i ekstrovertnosti i introvertnosti. Prema (Bobick & Johnson, 2001) od lakoće sa kojom se nadražuje moždana kora zavisi da li će neka osoba biti više ekstrovertna ili introvertna. Kod introvertne osobe određeni delovi mozga su veoma osetljivi na nadražaje i lako se stimulišu, zbog čega one preferiraju mirno okruženje i smirene situacije. Ekstrovertne osobe, sa druge strane, mogu da tolerišu visok nivo nadražaja moždane kore, zbog čega im je potrebno više socijalne interakcije i uzbudljivih dešavanja za stimulaciju.

2.4.5. Ostala istraživanja

U edukacionoj zajednici (Pianta, et al. 2008) interakcija između predavača i studenata se koristi kao metoda za procenu kvaliteta predavanja. Interakcija sama po sebi je kompleksan fenomen koji između ostalog uključuje sledeće aspekte: “emocionalni ton” u učionici i vezu između predavača i studenta; predavačevu sposobnost da odgovori na potrebe studenta; stepen interakcije predavača sa studentom; stepen koji predavač uspe da angažuje studenta i poveća njegovu sposobnost da uči tako što će sadržaj da prezentuje na interesantan način.

2.5. Potencijalni parametri

Rezultat pregleda multidisciplinarnih oblasti je veliki broj parametara za praćenje aktivnosti studenta. Proces praćenja aktivnosti odlikuje se karakterističnim zahtevima prema kojim je potrebno prilagoditi tehnike i način ekstrakcije parametara iz ambijenta. Potrebno je pre svega da se definišu parametri na osnovu kojih se može doneti zaključak o kvalitetu predavanja, razmotrile tehnike praćenja i obezbedilo prikupljanje istih pomoću odgovarajućih senzora i monitoring uređaja.

Ključni problem predstavlja izbor parametara koje je moguće prikupiti tako da se zadovolje sledeći zahtevi:

1. Merenje parametara bez postavljanja senzora na studente i u njihovoj neposrednoj blizini; na ovaj način se izbegava narušavanje sociološkog integriteta predavanja
2. Funkcionalna operativnost sistema u realnim (vanlaboratorijskim) uslovima

Praćenje aktivnosti upotrebom algoritama koji u van-laboratorijskim uslovima daju efikasne rezultate je prioritet da bi se obezbedio proces analize u realnom vremenu i realnim okolnostima. Sagledavanje kompleksnosti pojave upotrebe algoritama za praćenje određenih paterna detaljnije je objašnjena sledećim primerom: mehanizmi za prepoznavanje govora pružaju prihvatljivu tačnost kod izolovanog govora, ali imaju tendenciju rasta procenta greške u bučnim okruženjima. Performanse konvencionalnog mehanizma za prepoznavanje govora treniranog u izolovanom okruženju koji postiže stopostotnu tačnost u istom, opadaju do 30% u automobilu u pokretu (Lockwood & Boudy, 1992). Sa minimalnim prisustvom buke pri treniranju algoritama, greška se kreće i do 50% kada se prepoznavanje vrši u okruženju sa srednjim intezitetom buke (Bakis et al. 1993). Dodatan nivo buke utiče na signal menjajući vektore koji predstavljaju govor. Kao rezultat, dolazi do ključnih razlika između signala u trenažnom procesu i onih zabeleženih u realnom okruženju. Sistem treniran pod određenim SNR (eng. signal to noise ration), daje loše rezultate čak i kad se prepoznavanje reči vrši u povoljnijim SNR uslovima. U ovom slučaju cilj je da se trenažno okruženje što manje

razlikuje od okruženja u kome će se vršiti merenje (Gong, 1995). Prepoznat problem je aktuelan ne samo kod obrade zvuka već i za sve ostale procese ekstrakcije iz dinamičnog vanlaboratorijskog ambijenta.

2.5.1. Ostala istraživanja

U tabeli III predloženi su svi potencijalni parametri, dostupnosti metoda za njihovo prikupljanje kao i uticaj na sociološki integritet pri obradi audio/video signala na kome se nalazi veliki broj ljudi; u slučaju merenja i analize metodama primenljivim za grupna vanlaboratorijska merenja koja ne dolaze u sukob sa zahtevima; odnosno bez postavljanja senzora na studente i u njihovoj neposrednoj blizini kako bi se izbegao uticaj na sociološki integritet. Parametri čije prikupljanje ima uticaj na sociološki integritet nije efikasno meriti zbog tendencije delovanja na tačnost celekupnog sistema. Upotreba istih metoda u grupnim merenjima, na udaljenosti koja ne remeti kontinuitet i integritet predavanja sa zastupljenom pozadinskom bukom neće dati upotrebljive vrednosti.

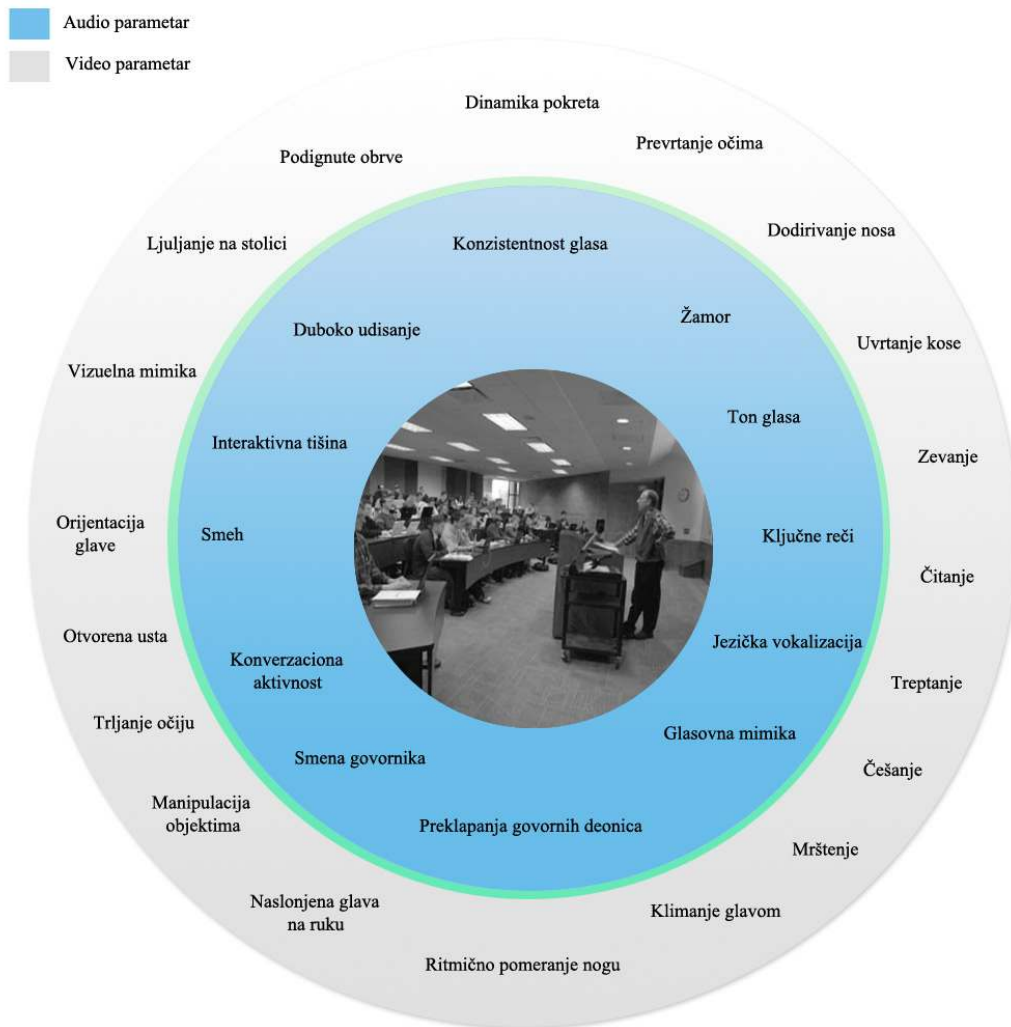
Kao što je moguće zaključiti pregledom tabele III, broj parametara i obrada svakog pojedinačno prevazilazi obim jedne studije. Shodno tome, presek parametara koji ispunjavaju zahteve i njihova korelacija je krucijalna, i ovi kriterijumi su preduslov za implementaciju sistema koji bi obezbedio zadovoljavajuću tačnost obrađenih rezultata. U drugoj koloni prikazano je da li se prikupljanjem parametra zadržava sociološki integritet. Treća kolona daje pregled dostupnosti metoda za prikupljanje i analizu pri grupnim merenjima sa distance koja ne remeti tok predavanja ali pruža zadovoljavajuću tačnost rezultata.

Tabela 4 Uticaj na sociološki integritet i dostupnost metoda za praćenje aktivnosti studenata

Parametar	Sociološki integritet	Dostupnost metode
Ključne reči	x	x
Ton glasa	x	√
Konzistentnost glasa	x	√

Dinamika pokreta	√	√
Smeh	√	√
Prevrtanje očima	x	√
Smena govornika	x	x
Orijentacija glave	x	√
Klimanje glavom	x	x
Zevanje	√	√
Konverzaciona aktivnost	√	√
Glasovna mimika	x	√
Vizuelna mimika	x	x
Mrštenje	x	x
Ljuljanje na stolici	x	x
Dodirivanje nosa	x	x
Uvrtanje kose	x	x
Manipulacija objektima	x	x
Interaktivna tišina	√	√
Preklapanja govornih deonica	√	√
Žamor	√	√
Glasovni segmenti	√	√
Govorna aktivnost	√	√
Stres govornika	√	√
Treptanje	x	√

Prema navednim zatevima (1) i (2) izvršeno je korigovanje liste parametara i ispitani su načini merenja i analize potencijalnih parametara. Parametri koji ne ispunjavaju zahteve nisu uključeni u dalju analizu.



Slika 7 Parametri za praćenje aktivnosti

Na slici 7 prikazani su svi parametri za detekciju aktivnosti koji se mogu pojaviti u svakodnevnom procesu predavanja. Dalji pregled i analiza, kao i metode koje se mogu koristiti za prikupljanje i obradu parametara, predmet je diskusije u nastavku teksta. Detaljnije su analizirani dinamika pokreta, nivo zvuka, smeh, orijentacija glave, konverzaciona aktivnost, žamor, glasovni segmenti, govorna aktivnost, stres i zevanje.

2.5.1.1. Deskriptori pokreta kao pokazatelj dinamike pokreta

Neverbalno ponašanje uključuje pokrete tela, statičku konfiguraciju tela bez pokreta, gestikulacije kao i dinamične pokrete delova tela: ruku, šaka, glave, delova lica, itd (Booch et al. 1999). Dinamika (intenzitet) pokreta (eng. fidgeting) predstavlja neposredan pokazatelj aktivnosti, koji u kombinaciji sa drugim parametrima može dati verodostojne rezultate i koristiti se za praćenje aktivnosti auditorijuma. Obzervaciona istraživanja pokazuju da je opadanje pažnje praćeno određenim promenama u ponašanju. Ponašanja poput vrpoljenja, zevanja, crtanja i gledanja okolo smatraju se indikatorima neaktivnosti (Varner & Dickinson, 1967).

Postoji mnogo načina da se izrazi određeno neverbalno ponašanje, zbog čega je ambiciozno meriti sve moguće vrste pokreta i klasifikovati sva njihova stanja. Prema tome, sami deskriptori pokreta globalno, mogu dati određenu informaciju o znatno kompleksnim kontekstima, zbog čega su oni upotrebljeni u ovom istraživanju kao pokazatelj dinamike pokreta. Dinamika pokreta u ovom slučaju predstavlja subjektivan stepen aktivnosti koji se može opaziti odnosno sam stepen pokreta (Peker et al. 2001). Ova metoda se obično upotrebljava u video kompresiji, kada se scena frejmovi bez vidljivih razlika kompresuju jer imaju malu stopu aktivnosti, dok se scene npr trke automobila ne mogu kompresovati jer imaju visoku stopu aktivnosti

U pregledu socioloških istraživanja, koje je dato u ovom radu, konstatovano je da se svaka bazična emocija sastoji od jednog procesa u mozgu koji se manifestuje setom komponenti: vokalnim izrazima, pokretima ruku, izraza lica, itd. Na osnovu ove činjenice, odrađene parametre je moguće indirektno meriti praćenjem deskriptora pokreta, jer oni predstalljaju sporadične pojave ovih komponenti. Nivo aktivnosti pokreta može se zabeležiti kamerom i pasivnim infracrvenim senzorom (PIR – eng. *Passive Infrared*) ili npr upotrebom senzora (akcelerometra) ili na neki treći način (npr ekstrakcijom pokreta sa slike).

2.5.1.2. Nivo zvuka

Povećan nivo buke može da prouzrokuje smanjenje koncentracije studenta izazivajući greške u interpretaciji predavanja. Evidentno je da povećavanjem nivoa buke opada pažnja studenta. Sa druge strane, ako je predavanje neinteresantno, studenti će praviti više buke. Da bi se izolovao nivo buke koji potiče od govora od buke drugog karaktera može se upotrebiti algoritam opisan u (Yamada et al. 2001). Nivo zvuka predstavlja nivo zvučnog pritiska u zatvorenom prostoru za čije merenje se koriste A-frekvencijska i F-vremenska ponderaciona kriva. Prema tabeli IV, nivo buke u učionici ne bi trebalo da prelazi vrednost od 60 dB. Vrednosti manje od 60 dB dalje će biti kategorizovane nakon izvršenih inicijalnih merenja. Parametar ispunjava zahteve i njegovo praćenje je izvodljivo.

Tabela 5 Jačina nivoa buke prema izvoru (Merenje buke, 2014)

Efekat	Jačina[dB]	Tipični izvor zvuka
Ozbiljno oštećenje sluha	140	Lansiranje rakete u blizini lansirne rampe
Oštećenje sluha i bol	130	Motor mlaznjaka (na rastojanju od 25m)
Oštećenje sluha posle kratkog izlaganja	120	Sirena za vazdušni napad (na rastojanju od 5m)
Ozbiljna opasnost od oštećenja sluha	110	Koncert tok muzike uz binu
Opasnost po sluh	100	Uzletanje mlaznjaka (300m)
Izvesna opasnost po sluh	90	Bučna industrijska hala
Dejstvo na zdravlje	80	Težak kamion, 70 km/h (na rastojanju od 10m)
Izvesna dejstva na zdravlje	70	Automobil, 70 km/h (na rastojanju od 10m)
	60	Normalan razgovor na rastojanju od 1m
	50	Tih razgovor na rastojanju od 1m
	40	Prigušena muzika sa radija
	30	Šaputanje na rastojanju od 1m
	20	Tiha spavaća soba

	10	Šuštanje lišća
Neprijatno "tiho"	0	Prag čujnosti

Tišina: Ako je izmeren nivo buke u učionici nizak, a nivo svih ostalih parametara koji reflektuju nezainteresovanost ili dosadu mali, može se zaključiti da je stepen pažnje visok.

2.5.1.3. Smeh

Kao što je već navedeno, smeh se detektuje primenom binarne klasifikacije poput metode potpornih vektora (eng. *Support Vector Machines* - SVM) i tehnike koje se često koriste kod prepoznavanja govora *Mel frequency cepstral coefficients* (MFCC). Druga metoda je perceptualno linearno predviđanje sa Gausovom metodom i neuralnim mrežama (Vinciarelli et al. 2009). Pomenuta istraživanja se odnose samo na audio signale, ali pojavile su se studije koja se bave i audio-vizuelnim prepoznavanjem nelingvističkih vokalnih signala.



Slika 8 Smeh u učionici

Detektor smeha kombinuje izlaz audio detektora koji koristi MFCC karakteristike govora i vizualnog detektora koji prepoznaje prostorne lokacije karakterističnih tačaka

lica (Ito et al. 2005). Detekcija smeha iz zapisa na kome se nalazi veliki broj ljudi je podložno greškama. Parametar ne ispunjava jedan od zahteva.

Smeh kao sociološki signal može se tumačiti dvojako kao odobravanje ali i u negativnom kontekstu zbog čega je neophodno sagledati njegovo značenje uporedo sa ostalim parametrima koji se mere da bi se razumeo kontekst u kom se ovaj parametar zabeležen u učionici.

2.5.1.4. Orijentacija glave

Orijentacija glave - je pouzdan način da se odredi u kom smeru je usmerena pažnja jedne osobe. Nekoliko studija je pokazalo da osobe podsvesno, gledaju u objekat ili uređaj sa kojim vrše interakciju. Ovaj parametar se može iskoristiti za analiziranje interakcije između studenata, kao i nivoa pažnje za studente pojedinačno (Stiefelhagen et al. 2006). Međutim, praćenje pravca pogleda grupe ljudi nije moguće realizovati jer zahteva pojedinačno snimanje svake osobe u grupi, odnosno upotrebu velikog broja kamera postavljenih u blizini studenata. Parametar ne ispunjava zahteve.

2.5.1.5. Konverzaciona aktivnost

Konverzaciona aktivnost – predstavlja preklapanje govornih deonica, koje se u svakodnevnom razgovoru kreće i do 10% (Shriberg et al. 2001) . Kao što je već navedeno kada obe strane aktivno učestvuju u konverzaciji, sagovornik prepoznaje momenat kada je došao red da preuzme/prepusti reč. U okruženju kakvo je učionica, preklapanje govornih deonica varira u zavisnosti od tipa predavanja. Basu et al. (Basu, 2002) je položio osnovu analize strukture konverzacije, pre čijeg rada nije bilo dovoljno detaljnih studija koja bi pomogla analizu konverzacije iz zvuka. Ovaj algoritam detektuje različite tipove konverzacije i predviđa promenu govornika. Dodatna vrednost algoritma je njegova mogućnost da segmentira zvuk na delova sa i bez govora i

robustan je po pitanju analize signala lošijeg kvaliteta i na distancu postavljenog mikrofona. Algoritam se može koristiti za analizu u segmentaciju zvuka u bučnom okruženju kod praćenja tona, za procenu govorne stope i energije, segmentiranje signala u kome se nalazi više govornika. U daljem tekstu ovaj algoritam referenciraće se kao Basu algoritam.

2.5.1.6. Žamor

Žamor je pozadinski šum koji se javlja kao rezultat preplitanja govora više osoba. Pretpostavka je da u procesu predavanja vrednost pozadinskog šuma ima malu vrednost ukoliko je zainteresovanost auditorijuma velika. Pri detekciji govornih segmenata povećan nivo žamora manifestuje se manjim brojem denoica koje se mogu detektovati kao govor, tako da se može reći da je žamor obrnuto proporcionalan broju detektovanih govornih segmenata.

2.5.1.7. Glasovni segmenti

Najjednostavniji sociološki signal je detektovan broj glasovnih segmenata (eng. *Voiced segments*) koji je inače u korelaciji sa zainteresovanošću (Dunbar, 1998) i ekstroverzijom (Nass & Brave, 2004). Ovaj parametar se koristi za modelovanje aktivnosti pri konverzaciji koja se dobija kao rezultat procentualnog vremena govora. Koliko neka osoba govori direktan je pokazatelj njegove dominacije u toku razgovora, a u kontekstu predavanja, dominantnost se može manifestovati kao znak sigurnosti u temu koja se prezentuje, što neposredno može da utiče na kvalitet prezentacije.



Slika 9 Glasovni segmenti u učionici

2.5.1.8. Stres

Prozodija je karakteristika govora dužeg od jednog fonetičkog segmenta i definiše se kao stres, akcentovanje i ritam (Werner & Keller, 1994). Stres u tonu može da bude nameran (govornik namerno naglašava deonice govora), ili nenameran (psihološki stres izazvan nelagodnošću). Vokalni stres može da predstavlja dinamičniji govor, gde treperenje (drhtanje) glasa može da označava psihološku reakciju zbog čega se tumači kao znak slabosti.

Stres u glasu se meri kao varijacija u intonaciji i naglasku (Handel, 1989) tako što se za svaki govorni segment ekstraktuje srednja vrednost energije, frekvencije i spektralne entropije. Stres iz signala je moguće izračunati tako što se prvo osnovni parametri ekstraktuju iz zvuka upotrebom Basu et al. algoritma, a potom se stres modeluje kao rezultat zbira standardne devijacije formantne frekvencije, spektralne entropije i energije u frejmu (Caneel, 2005).

Stres se takođe može ekstrahovati iz elektro-encefalogramskog signala upotrebom relativnog odnosa energije (eng. *Relative Energy Ratio* - RER), šenonskom entropijom (eng. *Shannon Entropy* - SE) i spektralnih centroida (eng. *Spectral Centroids* - SC) (Sulaiman et al. 2012). Relativni odnos energije spektralne gustine (eng. *Spectral Density* - SD) računa se za svaki frekvencijski opseg za četiri različite grupe, nakon čega se šenonska entropija upotrebljava za potvrđivanje paterna stresa. Rezultati simulacije prepoznavanja stresa gde je input za neuronske mreže korišćen skup opisanih parametara, sa setom za klasifikator 50:50 i 70:30 bili su 88.89% tačnosti.



Slika 10 Stres u učionici

2.5.1.9. Zevanje

Zevanje je pojava koja se može manifestovati zbog više razloga, među kojima su neki od interesa za ovo istraživanje. Tranzicija dana u noć može da izazove zevanje kod ljudi, ali jedna bitnija činjenica ukazuje na to da se zevanje može koristiti za detektovanje zainteresovanosti u učionici. Veća frekvencija zevanja povezana je sa gledanje neinteresantnog i repetativnog sadržaja.

Takođe, pokazano je da je zevanje studenata najčešće neposredno jedan sat pre spavanja kao i jutarnjeg buđenja. Dnevni prelazak iz stanja neaktivnosti u aktivno stanje (rano ujutru) i obrnuto, imaju jaku korelaciju sa zevanjem. Postojeća istraživanja ne podržavaju činjenicu da je zevanje povezano sa nedostatkom sna (Baenninger, 1997).

2.5.1.10. Analiza metoda ekstrakcije parametara

U cilju praćenja aktivnosti studenata, sledeći osnovni parametri su izdvojeni: nivo zvuka, dinamika pokreta, smeh, žamor, glasovni segmenti, govorna aktivnost i stres (Tabela 4). Dalje istraživanje biće usmereno na rešavanje problema praćenja aktivnosti studenata i pronalaženja metoda koje bi iskoristile dostupne parametre koje ispunjavaju zahteve i obezbeđuju obradu podataka u realnom vremenu.

Tabela 6 Parametri za praćenje aktivnosti

Parametar	Senzor	Ispunjava zahteve	Tehnika
Dinamika pokreta	Kamera/PIR	√	komparacija razlike piksela, strukturna razlika između objekata DSSIM, MHI (<i>Motion History Images</i>)
Nivo zvuka	Mikrofon	√	A-frekvencijska ponderaciona kriva
Smeh	Mikrofon	x	C4.5 klasifikator
Orijentacija glave	Kamera	x	3D procena pozicije
Tišina	Mikrofon	√	A-frekvencijska ponderaciona kriva
Konverzaciona aktivnost	Mikrofon	√	Basu algoritam
Žamor	Mikrofon	√	A-frekvencijska ponderaciona kriva
Vreme govora	Mikrofon	√	Basu algoritam
Govorna aktivnost	Mikrofon	√	Basu algoritam
Stres	Mikrofon	√	Basu algoritam
Konverzaciona aktivnost	Mikrofon	√	Izveden od više različitih metoda

Zevanje	Kamera	√	Kalmanov filter i srednja vrednost promene aktivne IR iluminacije
---------	--------	---	-------------------------------------------------------------------

Kao rezultat praćenja aktivnosti studenata pojavljuje se skup različitih parametara čiji zaseban pregled ne daje mnogo informacija o zainteresovanosti i stepenu pažnje. Prikupljene podatke je potrebno predstaviti određenim modelom koji definiše korelaciju različitih parametara; na primer, nezainteresovanost se može posmatrati kao povećan nivo buke i pokreta u datom momentu.

Visok nivo buke i visok nivo pokreta nedvosmisleno ukazuju na postojanje nezainteresovanosti ili nekog oblika sukoba ili rasprave u auditorijumu. Dodatan žamor potvrđuje pretpostavku da je nivo aktivnosti nizak. Kada nivo zvuka ne prevazilazi jačinu normalnog razgovora pretpostavka je da se u učionici predavanje odvija povoljno ili su studenti pasivno nezainteresovani. Dati modeli trenutno nisu dovoljno deskriptivni da se iz njih može zaključiti da li su studenti pasivno-zainteresovani ili nezainteresovani.

Takođe, predloženi parametri su previše kompleksni da bi se mogli ekstrahovati na jednostavan način upotrebom jedne metode, zbog čega je potrebno upotrebiti mnogo različitih tehnika (u ovom slučaju 4 različite tehnike za 9 parametara iz zvuka, što može da bude dosta zahtevno po pitanju resursa). Prema tome, da bi se bolje razumeo način prikupljanja podataka kao i da bi se optimizovala metoda ekstrakcije i detekcije parametara (da bi se parametri mogli detektovati na jedinstven način upotrebom manjeg broja algoritama i IoT uređaja; na primer, većina parametara koji se ekstrahuju iz zvuka imaju određene karakteristike zvuka koje je potrebno detektovati a imaju ih i drugi parametri i mogu se zabeležiti istim IoT nodom).

Tabela 7 Struktura parametra koji odgovaraju zahtevima i spisak karakteristika koje se koriste za njihovu konstrukciju

Parametar	Osnovni parametar	Metoda ekstrakcije	Algoritam	Komentar
Dinamika pokreta	Frekvencija pokreta	Računanje strukturalnih razlika između različitih frejmova	DSSIM	Obrnuto proporcionalno SSIM
Nivo zvuka	Formantna frekvencija, autokorelacija	Merenje visine formantata i najvećeg pika autokorelacije	Basu	Umesto nivoa zvuka predlaže se merenje formantne frekvencije i autokorelacije koje se ekstraktuju kao sastavni deo Basu algoritma.
Tišina	Segmenti bez govora i detektovanog zvuka i malim vrednostima autokorelacije	Obrnuto proporcionalan broj segmenata sa detektovanim govorom i detektovanog zvuka sa malim vrednostima autokorelacije	Basu	Nije potrebno koristiti dodatan algoritam, već se tišina može detektovati kao interaktivna tišina (tišina nasuprot dinamičnim govornim segmentima). Segmenti sa malim vrednostima autokorelacije i bez detektovanog govora

Konverzična aktivnost	z-score govora i segmenata	vremena govornih	Segmentacija signala na delova sa i bez govora. Ekstrakcija govorne aktivnosti i vremena govora. Konverzična aktivnost se dobija kao z-rezultat zbira govorne aktivnosti i vremena govora.	Basu	Izveden parametar koji se dobija naknadim izvođenjem korišćenjem više osnovnih parametara
Žamor			Broj deonica koji nisu identifikovani kao govor, mali autokorelacioni vrhovi	Basu	Neophodno je dodatni proširiti basu algoritam da biautomatski radio ekstrakciju i žamora
Glasovni segmenti			Dobija se segmentacijom signala na delove sa i bez govora upotrebom kvadratne sredine (RMS)	Basu	
Govorna aktivnost			Segmentacija signala na delova sa i bez govora.	Basu	

Stres	z-score standardne formantne frekvencije, spektralne energije u frejmu	zbir devijacije entropije i	Segmentacija signala na delova sa i bez govora. Ekstrakcija govorne aktivnosti i vremena govora. Konverzaciona aktivnost se dobija kao z-rezultat zbira govorne aktivnosti i vremena govora.	Basu	Izveden parametar koji se dobija naknadim izvođenjem korišćenjem više osnovnih parametara
-------	------------------------------------------------------------------------	-----------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------	-------------------------------------------------------------------------------------------

Pregledom datih osnovnih parametara za prethodno navedene složene parametre, stvoren je uslov da se uvrsti još jedna dodatna karakteristika iz zvuka koja se modeluje korišćenjem već postojećih predloženih – stres: dobija se modelovanjem z-score vrednosti zbira standardne devijacije formantne frekvencije, spektralne entropije i energije u frejmu.

2.5.1.10.1. Analiza hardverskih i softverskih rešenja

Od navedenih hardverskih jedinica mogu se upotrebiti mikrofon da bi se omogućilo snimanje zvuka, kamera za ekstrakciju pokreta i pasivni infrared senzor (PIR) koji bi takođe detektovao pomeranje u auditorijumu. Da bi se izbegla redundantna upotreba uređaja, urađena je analiza svake hardverske jedinice ponaosob i na kraju je dat finalni zaključak na osnovu potreba sistema i zahteva algoritama za ekstrakciju parametara iz digitalnog signala. U daljem istraživanju biće uvršten niz dodatnih parametara za prepoznavanje pažnje nakon obzervacione faze istraživanja, kako bi se približnije mogao odrediti karakter datog modela na osnovu varijabli dobijenih iz realnog okruženja. Metode mašinskog učenja se mogu primeniti kako bi se izvršilo prepoznavanje parametara u okruženju. Da bi se ovo ostvarilo potrebno je prvo prikupiti uzorak koji bi služio za treniranje. Finalna lista parametara predstavljena je u jednom od narednih poglavlja gde su između ostalog, ekstraktovani i osnovni parametri koji se koriste za modelovanje gore navedenih: formantna frekvencija, spektralna entropija, autokorelacija i stopa glasova.

Mikrofon

Mikrofon predstavlja akoustično-električni provodnik ili senzor koji pretvara zvuk u električni signal. Mikrofonu su korišćeni u mnogim aplikacijama a zavisno od primene postoje mikrofoni koji koriste elektromagnetnu indukciju, promenu kapaciteta, ili modulaciju da bi se dobila električna voltaža iz mehaničke vibracije.



Slika 11 Mikrofon

Za ovo istraživanje moguća je upotreba sledećih pristupa za snimanje zvuka:

- veći broj mikrofona koji bi se implementirali u određeni deo učionice tako da se ravnomerno pokrije cela učionica;
- upotreba pojedinačnih mikrofona za svakog učesnika u predavanju ponaosob (mikrofoni se postavljaju na studenta u blizini usta studenta/profesora)
- upotreba jednog mikrofona.

Ako se koristi veći broj mikrofona, dodatno se komplikuje implementacija, kao i ekstrakcija zbog neophodne sinhronizacije većeg broja izvora kao i kasnije analize svakog izvora pojedinačno. U suprotnom, kad se zvuk snima upotrebom jedne senzorske jedinice snima se i ambijentalni eho u prostoriji kao i prirodna mešavina zvukova. Ovo može da predstavlja problem ako algoritam nije robustan na veći broj govornika u datom digitalnom zapisu, kao i zbog šumova.

Kamera

Kako bi se uradila analiza strukturalnih razlika iz frejmova, potrebna je kamera ili digitalan aparat koji je moguće programirati da u određenom intervalu pravi sliku ekrana (eng. Screenshot) sa mogućnošću pristupa snimljenim frejmovima. Većina danas dostupnih kamera istovremeno snima i zvuk, tako da je mikrofoni moguće eliminisati u krajnjoj implemetaciji i koristiti mikrofoni koji je deo video seta za snimanje.



Slika 12 Kamera

Kvalitet slike i zvuka je bitan ali ne i presudan faktor. Kako je navedeno ranije potrebno je da trenažno okruženje bude što približnije okruženju u kome će se vršiti prepoznavanje i detekcija parametara, tako da je zadovoljavajući kvalitet slike i 1024x768 dok je algoritam koji se koristi za ekstrakciju robustan kvaliteta slike.

Trenutno, pametni mobilni telefoni imaju kvalitetne objektivne i veliki broj megapiksela, zavidne procesorske i memorijske resurse, mikrofona za snimanje zvuka, kao i mogućnost razvoja velikog broja aplikacija upotrebom Android platforme. Navedeni razlozi opravdavaju izbor Android mobilnog uređaja da se u ovom istraživanju koristi kao instrument za prikupljanje podataka. U nastavku detaljnije je objašnjen uređaj koji je korišćen u istraživanju.

Android telefon

U ovoj studiji korišćen je Android Samsung Galaxy SII, u momentu pisanja ovog rada jedan od Android telefona sa kvalitetnijom kamerom od 8MP (Mega Piksela) i mikrofonom.

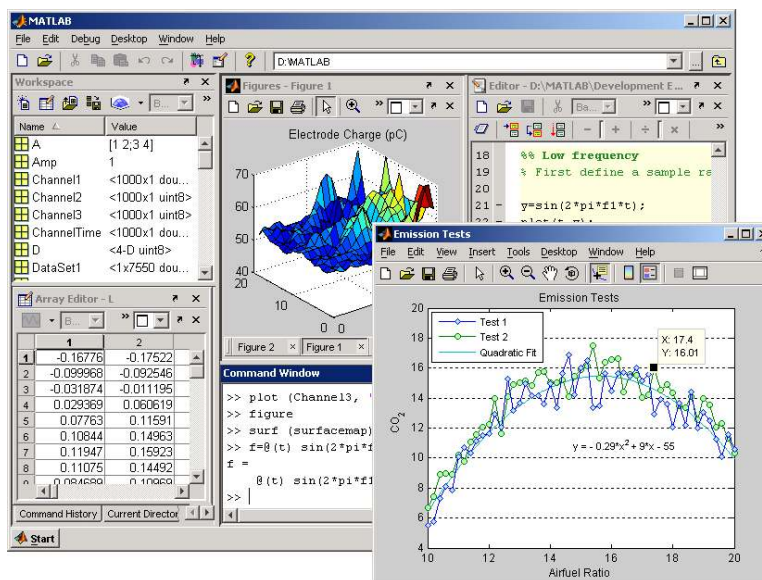


Slika 13 Android pametni telefon korišćen u istraživanju

Jedina mana ovog uređaja predstavlja mikروفon koji je dizajniran za prenos govora sa bliske distance, međutim ovo je prevaziđeno tako što je ekstrakcija i procesiranje zvuka urađena pomoću algoritma koji je robustan na kvaliteta zvuka, broja govornika i intezitet šuma prisutnog na zapisu.

Matlab

Matlab je moćan softverski alat za procesiranje zvuka, klasifikaciju i prepoznavanje paterna, sa mnoštvom alata, kompatibilnih sa ogromnim brojem standarda i tehnologija. Karakteristike ovog softvera čine ga pogodnim alatom za sprovođenje datog istraživanja, jer se praktično sve neophodne funkcionalnosti mogu razviti preko različitih Toolbox-ova: procesiranje signala, treniranje i prepoznavanje paterna upotrebom metoda mašinskog učenja, pa čak i sama prezentacija obrađenih rezultata. Prednost upotrebe Matlaba u odnosu na druge platforme je taj što nije potrebno razviti sve funkcionalnosti sistema, jer su one već dostupne i gotovo momentalno se mogu povezati u jednu celinu upotrebom Toolbox-ova; dok sa druge strane za Androida platformu neophodno je razviti funkcionalnosti koje nedostaju.



Slika 14 Matlab softverski alat

Uporedni pregled platformi za razvoj sistema

Tabela 8 Uporedni pregled funkcionalnosti platformi za razvoj sistema: Matlab i Android

	Matlab	Android
Procesiranje slike	Image Processing Toolbox	OpenCV (OpenCV, 2013)
Procesiranje zvuka	Data processing Toolbox	jMIR2.0 *
Klasifikacija	Neural Network Toolbox	Weka** (Weka Android, 2013), (Hall et al. 2009)
Prepoznavanje parametara	Neural Network Toolbox	Weka** (Weka Android, 2013), (Hall et al. 2009)
Vizualizacija	Data Visualization Toolbox	JUNG (JUNG, 2013) i mnoge druge <i>custom</i> biblioteke
Izvršavanje simultanih radnih procesa	Parallel Computing Toolbox	AsyncTask
Basu algoritam	Kompletna implementacija u Matlabu	Ne postoji u Javi
Bajesove mreže	Bayes Network Toolbox	Weka*, bayes-android**, (Bayes Network Android, 2013) Bayesian Network Classifier* (Baesian Network Java, 2013)

*Biblioteka u Javi koju je potrebno portovati na Android

**Bibiloteka postoji za Android ali je nedovoljno testirana

Pregled funkcionalnosti alata koji se mogu upotrebiti za razvoj platforme ukazuje da se kompletan sistem može razviti upotrebom Matlab. U slučaju da se koristi Android, potrebno je spojiti više biblioteka za svaki od procesa, dok je za neke biblioteke čak neophodno portovati kompletnu biblioteku na Android ili je biblioteka nedovoljno testirana da bi mogla da se upotrebi u projektu. Bajesove mreže su neophodne u sistemu jer ih Basu algoritam koristi za segmentaciju signala. U Matlabu postoji već gotov Toolbox (Bayes Network Matlab, 2007), ali sama Android platforma nema zadovoljavajuću kompletnu implementaciju Bajesovih mreža koja bi zadovoljila sve potrebe.

Što se tiče simultanih radnih procesa istovremeno obe platforme podržavaju multitasking: Matlab koristi Toolbox za paralelno izvršavanje procesa koji su definisani kao poslovi (eng. jobs); Android koristi asinhronne taskove koji omogućuju multithreading programiranje. Android klasa AsyncTask omogućava jednostavno simultano procesiranje zahteva, pozadinske operacije i objavljivanje rezultata bez potrebe za manipulacijom UI tredovima i/ili hendlerima.

Potreba da se istovremeno snime i obrade zvuk i slika tako da isti budu vremenski sinhronizovani upotrebom jednog od dve prethodno navedene platforme dodatno otežava proces razvoja kod Matlab-a. Pored toga što Matlab omogućava multitasking za ogroman broj procedura, samo prikupljanje digitalnog signala upotrebom različitih Toolbox-ova nije uvek moguć. Ovaj problem se može prevazići pokretanjem dve instance Matlab-a od kojih se jedna može upotrebiti za snimanje zvuka a druga za snimanje slike.

2.6. Zaključak

Pregledom oblasti kroz jedinstvenu kategorizaciju koja je data u ovom radu, dobijen je kompletan pregled i uvid u stepen trenutnog razvoja inteligentnih učionica; uočava se preplitanje i intezitet kojim se različite tehnologije dopunjuju, obrazujući kompleksne platforme nastale kao rezultat implementacije različitih tehnologija pametne učionice u sopstvenoj režiji.

Dokazano je da pametne učionice podižu efektivnost i pozitivno se odražavaju na akademski učinak studenata. Obzirom na to da trenutna istraživanje iz ove oblasti nude dosta prostora za napredovanjem i da ni jedno rešene nema mogućnost praćenja aktivnosti studenata; može se zaključiti da sistemi koji nude uvid u nedostatak pažnje i pregled faktora koji je reflektuju u realnom vremenu u toku predavanja predstavljaju naučni doprinos, kao i da mogu da ostvare bolji rezultat od postojećih rešenja.

Navedeni sociološki/psihološki faktori upotrebljivi su za jasnije razumevanje samog porekla primećene aktivnosti u auditorijumu. Možda deo navedenih pokazatelja ne deluje dovoljno relevantno (poput ekstroverzije), međutim, činjenica je da razumevanje ljudske aktivnosti, zbog svoje kompleksnosti zahteva analizu velikog broja faktora. Neke od njih nije moguće izmeriti postojećom tehnologijom ili u realnom vremenu za vreme predavanja, ali sa tendencijom rasta sveopšte povezanosti ljudi, uređaja i Interneta, u skorijoj budućnosti pomoću IoT-a obezbediće se razumevanje konteksta koji do sada nisu bili u mogućnosti da se definišu.

Analiza potencijalnih parametara je data da bi se stekao uvid u parametre koje je moguće meriti i mogućnosti njihovog praćenja u realnom okruženju. Može se zaključiti da se većina karakteristika iz zvuka koje obrazuju ljudsko ponašanje prvenstveno ekstraktuju upotrebom većeg broja osnovnih parametara, koji se kasnije modeluju i čiji input služi za treniranje mašinskog algoritma koji je sposoban da identifikuje određene parametre. U daljem tekstu, kroz obzervaciono istraživanje validirani su predloženi parametri i napravljena je konačna analiza istih koji se efikasno mogu koristiti za prepoznavanje aktivnosti studenata u realnim uslovima. Predloženi su hardver i softver koji će se upotrebljavati za prikupljanje i ekstrakciju podataka.

Takođe, evidentno je da većina dosadašnjih istraživanja baziraju svoje rezultate na istraživanjima sprovedenim u laboratorijskim uslovima, odnosno informacije su prikupljene sensorima koji su postavljeni na osobe, zvuk je sniman samo za jednu osobu a kasnije analiziran, i sve to u uslovima koji su daleko od realnog okruženja. Predlog, kao i sama simulacija ove platforme omogućiće da se u realnim okolnostima odrede kognitivna grupna stanja, odnosno da se modeluje aktivnost, pažnja i zainteresovanost, koja do sada nisu mogla da se identifikuju. Pored toga, ovo je prvo istraživanje ovakvog tipa gde se upotrebom tehnologije dizajnira sistem za praćenje kvaliteta predavanja.

Pregledom tehnologija koje se mogu upotrebiti za razvoj sistema pokazalo se da je kompletan sistem moguće razviti u Matlab-u i da je proces razvoja komponenti lakši nego upotrebom drugih platformi.

3. Metodologija istraživanja

Proces razvoja sistema za prepoznavanje paterna je nemoguće realizovati bez klasifikatora, odnosno bez upotrebe neke od dostupnih metoda klasifikacije paterna poput metode potpornih vektora, i neuronskih mreža, Bajesovih mreža i drugih algoritama za mašinsko učenje. Takođe, evaluacija performansi sistema kroz simulaciju, obavljena je testiranjem tačnosti sistema na različitim skupovima podataka koja nisu korišćeni u procesu treniranja i validacije. Tačnost različitih metoda mašinskog učenja se ne može predvideti kod kompleksnih sistema, zbog čega je urađena uporedna klasifikacija za veliki broj metoda mašinskog učenja, kako bi se uradila selekcija.

Glavna metrika za evaluacije klasifikatora u ovom istraživanju su tačnost klasifikatora, broj tačnih pozitivnih klasifikacija (eng. True Positive), broj netačnih pozitivnih klasifikacija (eng. False Positive), preciznost (Witten et al. 1999) i kapa koeficijent.

Generalna šema za klasifikaciju prikazana uključuje: treniranje klasifikatora korišćenjem uzorka ekstraktovanog sa audio, video zapisa i akcelerometra, klasifikaciju i prepoznavanje paterna.

3.2. Ekstrakcija parametara

Ekstrakcija parametara izvedena je nad prethodno označenim deonicama od 30 sekundi, koje su manuelno izdvojene iz snimljenog materijala i klasifikovane kao interesantne i neinteresantne na osnovu ocena koji su studenti dali za vreme predavanja. Ekstraktovan je veliki broj parametara iz zvuka, deskriptori pokreta studenata sa slike kao i deskriptori pokreta predavača. Galvanska provodnost kože je u finalnoj verziji sistema odbačena kao parametar jer ne izlazi u susret zahtevima sistema. Eksperimentalni rezultati sa galvanskom provodnosti su uključeni kao dodatak u ovom poglavlju.

3.2.1. Ekstrakcija parametara iz zvuka

Za klasifikaciju određenih deonica iz zvuka ključno je da se prepoznavanje uradi što tačnije: da se razlikuju deonica u kojima se nalazi govor i u kojima nema govora, u kojima ima glasova i nema glasova, kako bi se dalje mogle izračunati govorna aktivnost, glasovna stopa, formantna frekvenkcija, itd. Neophodno je da algoritmi za prepoznavanje zvuka budu robusni na distancu postavljanja mikrofona i buke iz okruženja. U ovom radu korišćen je Basu algoritam koji omogućava efikasno preprocesiranja zvučnog signala u takvim uslovima kao i segmentaciju zvuka i prepoznavanje govornih intervala.

Zvuk je procesiran tako da su ekstraktovane sledeće karakteristike upotrebom Bayes Network Toolbox-a za Matlab: glasovni segmenti, govorna aktivnost, formantna frekvencija, spektralna entropija, autokorelacija i glasovna stopa (brzina izgovaranja glasova), kao i standardne devijacije i srednje vrednosti vremena govora, govorne aktivnost, formantna frekvencija, spektralna entropija i autokorelacije.

Radi detaljnije analize, iz prethodnih karakteristika zvuka izračunate su sledeće izvedene karakteristike: konverzaciona aktivnost i stres. Konverzaciona aktivnost se dobija kao z rezultat zbira vremena govora i govornih segmenta. Stres predstavlja zbir z rezultata standardne devijacije formantne frekvencije, spektrealne entropije i energije.

Prezhodna istraživanja pokazala su da su prethodno navedene karakteristike iz zvučnog signala u vezi sa izražavanjem ljudskog ponašanja koje se manifestuje kroz govor (Vinciarelli et al. 2009) (Pentland, 2004) (Vinciarelli et al. 2009). Analizirane su sekvence od 30 sekundi jel je u preliminarnom procesu klasifikacije postignuta veća tačnost nego kod sekvenci u trajanju od jedne minute.

Kao što je ranije navedeno interakcija u učionici se može identifikovati upotrebom sledećih karakteristika zvuka konverzacione aktivnosti, uticaja i stresa. Ove karakteristike se računaju upotrebom zvučnih parametara: formantna frekvencija, spektralna entropija, autokorelacija, energija, glasovni segmenti, vreme govora i glasovna stopa. U daljem tekstu objašnjena je ekstrakcija navedenih parametara.

3.2.1.1. Konverzaciona aktivnost

Aktivnost je vreme koja jedna osoba govori. Procenat govornog vremena koristi se kao parametar i predstavlja ukupnu visinu tona bez obzira na prekide i preklapanja govora od strane drugih. Procenat govornog vremena je u korelaciji sa nivoom zainteresovanosti (Dunbar, 2004) i računa se kao s/n gde je s broj govornih frejmova a n je ukupan broj frejmova u govornom segmentu. Frekvencija govora se meri kao stopa govora u govornim regijama konverzacione (Stoltzman, 2006):

$$v(v + u)$$

Gde su v frejmovi sa glasovima a u frejmovi bez glasova.

3.2.1.2. Uticaj

Uticaj koji jedna osoba ima na drugi za vreme interakcije se može modelovati kao Markovljev proces.

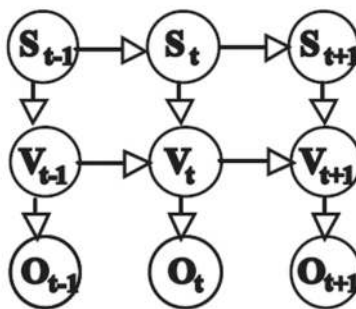
3.2.1.2. Stres

Stres se računa kao srednja vrednost standardne devijacije energije, formantne frekvencije i spektralne entropije:

$$\Sigma_{std}(\epsilon) + \Sigma_{std}(\mu) + \Sigma_{std}(\rho)$$

3.2.1.3. Ekstrakcija

Kao prvi korak, snimljeni zvuk se segmentira i ovi segmenti su dalje klasifikovani kao govorni i segmenti bez govora. Segmentacija se radi upotrebom skrivenog Markovljevog modela (Saul et al, 1995) prikazanog na slici ispod. Donji nivo sa slike (O_t) predstavlja kontinuiranu obradu parametara (autokorelacije, spektralne entropije, formantne frekvencije). Drugi nivo predstavlja govorna stanja (V_t); dok treći nivo predstavlja glasovna stanja (S_t).



Slika 15 Skriveni Markovljev model

Autokorelacija, broj autokorelacionih pikova i normalizovana spektralna entropija se računaju za svaki frejm iz govornih segmenata sa dužinom frejma od 32ms i preklapanjima od 16 ms između frejmova.

Autokorelacija signala $s[n]$ dužine N se ekstrahuje:

$$A[k] = \frac{\sum_{n=k}^N s[n]s[n-k]}{\left(\sum_{n=0}^{N-k} s[n]^2\right)^{\frac{1}{2}} \left(\sum_{n=k}^N s[n]^2\right)^{\frac{1}{2}}}$$

Spektralna entropija se koristi za detekciju deonice sa govorom iz zvuka kod prepoznavanja aktivnosti govornika. Snimljeni zvuk je u praksi često lošijeg kvaliteta entropija omogućava pouzdanije informacije o govornim i glasovnim segmentima. Da vi se izračunala spektralna entropija visina tona se normalizuje $P(w) = \frac{P(w)}{\sum P(w)}$ da bi se

otklonio uticaj energije u signalu. Konačno, relativna spektralna entropija se računa kao:

$$Hr = - \sum_w \left(p(w) \log \frac{p(w)}{m(w)} \right)$$

Formantna frekvencija je frekvencija ljudskog govora i ekstraktuje se na sledeći način:

$$F = \frac{fs}{2\pi} \text{ Hz}$$

Gde je fs frekvencija snimljenog zvuka od 8 kHz.

Kako bi se nivelisao uticaj energije u zvuku na ostale parametre, što bi zbačilo da je algoritam robustan na različite mikrofone i buku; dužina glasovnih segmenata i segmenata bez glasova je uprosečena za period od 30 sekundi. Energija (visina tona) signala se može izračunati kao koren kvadrata amplitude:

$$x_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

Gde je n broj smplova i x_i je signal. Energija frejma u vremeni se računa kao logaritam energije frejma. Konačno, kako bi se signal klasifikovao na govor i deonice bez govora, spektralna entropija i formantna frekvencija su korišćenje kao input u Markov model u dva navrata: prvi put se uradi segmentacija govora i rezultujući delovi signala bey govora se koriste da bi se odredila varijansa signala za vreme tišine.

Potom, izlaz Gausovog postupka upotrebom ove varijanse je dodat celokupom signalu i segmentacija je urađena ponovo.

Sledeći parametri iz zvuka su izvedeni iz prethodno ekstraktovanih parametara:

- Dužina glasovnih segmenata kao dužina samoglasnika

- Dužina govornih segmenata kao dužina fraze izračunata upotrebom glasovnih i govornih deonica

3.2.2. Ekstrakcija parametara sa slike

Svaka slika se sastoji od velikog broja informacija čije karakteristike obrazuju strukturu objekata na sceni. Osvetljenost površine objekata na sceni je proizvod odbijanja svetlosnih zraka, međutim sama struktura objekata je nezavisna od osvetljenja. Prema tome, da bi se istražila strukturalna informacija objekata, potrebno je izolovati uticaj osvetljenosti. Iz razloga što osvetljenje i kontrast mogu da variraju u sekvenci, lokalno osvetljenje i kontrast se mogu iskoristiti pri modelovanju razlike između objekata. Upotreba SSIM algoritma za praćenje pokreta objekata predmet je istraživanja (Loza et al. 2009), gde je pokazano da je moguće ovaj algoritam koristiti za praćenje pokreta. U ovom istraživanju, predložen algoritam je korišćen za određivanje strukturalne razlike između sekvenci.

SSIM index je mera vizuelno vidljivih razlika između dve slike i bazira se na poređenju osvetljenja $l(f,g)$, kontrasta $c(f,g)$ i strukture $s(f,g)$:

$$SSIM = [l(f,g)^\alpha] [c(f,g)^\beta] [s(f,g)^\gamma]$$

α , β i γ variraju zavisno od važnosti svake od komponenti u određenoj primeni algoritma. Index strukturalna razlike se direktno izvodi is SSIM-a na sledeći način:

SSIM koristi strukturalne informacije sa slike kao attribute koji predstavljaju strukturu objekata na sceni. Ovaj algoritam je robustan na promene osvetljenja i kontrasta. DSSIM je izmeren na frejmovima zabeleženi u vremenskom intervalu od 30 sekundi i ta vrednost je u daljem tekstu deskriptor pokreta studenata u datom periodu.

Frekvencija pokreta studenata ekstraktovana je u svakoj sekundi između trenutnog i narednog frejma upotrebom indeksa strukturalne razlike (eng. Structural dissimilarity - DSSIM), pokazatelja obrnuto proporcionalnog indeksu strukturalne sličnosti (eng. Structural similarity - SSIM) (Wang et al. 2004):

$$DSSIM(x, y) = \frac{1}{1 - SSIM(x, y)}$$

gde su x i y prethodni i naredni frejm.

Sistemski dijagram SSIM algoritma prikazan je na slici 16. Sličnost između frejmova se meri poređenjem tri karakteristike: osvetljenja, kontrasta i strukture. Prvo se osvetljenje svakog signala uporedi kao mera srednjeg inteziteta:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i$$

Funkcija za upoređivanje osvetljenja $l(x,y)$ je μ_x i μ_y . Nakon toga se odstrani intezitet iz signala. Rezultujući signal $x - \mu_x$ odgovara projekciji vektora x na hiperravni definisanoj kao:

$$\sum_{i=1}^N x_i = 0$$

Za procenu kontrasta signala koristi se standardna devijacija (kvadratni koren varijanse), čija se procena računa kao:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}}$$

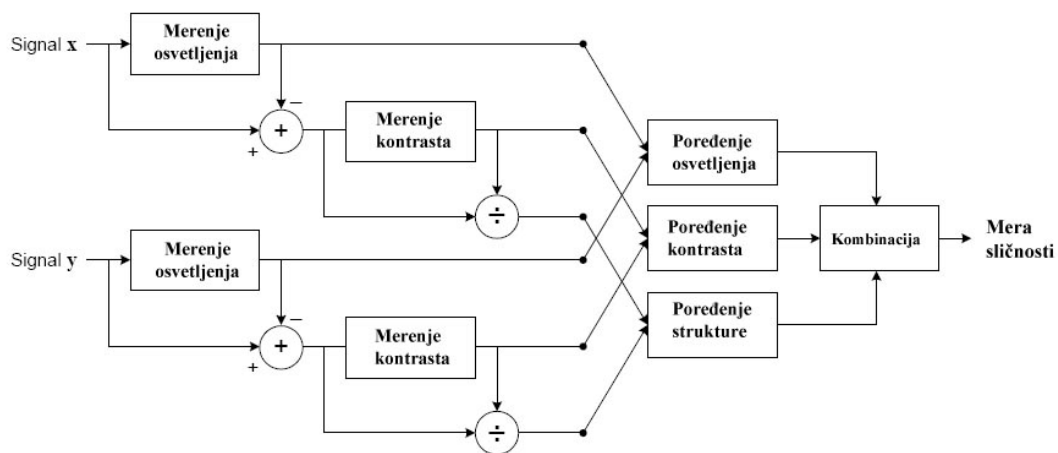
Poređenje kontrasta $c(x;y)$ se dalje računa kao σ_x i σ_y . Nakon prethodnih procesiranja, signal se normalizuje prema sopstvenoj standardnoj devijaciji. Strukturno poređenje $s(x;y)$ se sprovodi na prethodno normalizovanim vrednostima $(x - \mu_x) / \sigma_x$ i $(y - \mu_y) / \sigma_y$.

Konačno, tri komponente se kombinuju da bi se dobila mera sličnosti:

$$S(x; y) = f(l(x; y); c(x; y); s(x; y))$$

Značajno je napomenuti da su sve tri komponente relativno nezavisne. Promena osvetljenosti i/ili kontrasta neće uticati na samu strukturu slike. Da bi se kompletirala definicija indeksa sličnosti, definisane su tri funkcije $l(x; y)$, $c(x; y)$, $s(x; y)$, kao i kombinacije ovih funkcija.

SSIM algoritam koristi strukturne informacije sa slike kao atribute koji predstavljaju strukturne objekte na sceni. Ovaj algoritam je robustan na promene kontrasta i osvetljenja koji se često mogu pojaviti za vreme snimanja i uticati na finalni rezultat.



Slika 16 Dijagram algoritma za merenje strukturalne sličnosti

Srednja vrednost indeksa strukturne razlike između frejmova zabeleženih kamerom su upoređene da bi se dobio ukupan broj promena u vremenu od 30 sekundi, koji

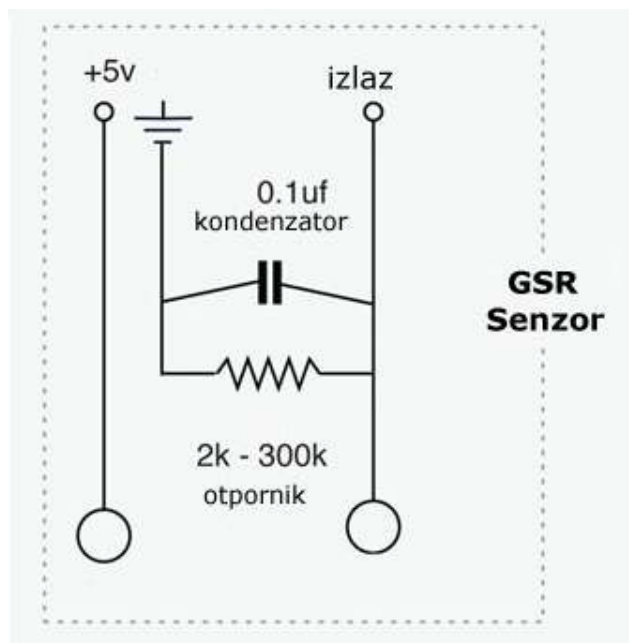
predstavlja razliku između objekata na sceni (Ulrici et al. 2010) i frekvenciju pokreta u datoj sekvenci.

3.2.3. Merenje provodnosti kože

Galvanska provodnost kože je jedna od nekoliko elektrodermalnih reakcija koje predstavljaju promene u električnim karakteristikama kože spoljašnjim uticajem na pojedinca, čime se može dovesti uticaj okruženja na psihološko stanje osobe. Ljudska koža je dobar provodnik elektriciteta i kada se slaba električna struja sprovede kroz kožu, promene u provodnosti kože se mogu izmeriti (Know & Holden, 1963). Varijable koje se mere su ili otpornost kože ili njegova recipročna vrednost - provodnost kože. Prema Omovom zakonu otpornost (R) jednaka je voltaži (V) između dve elektrode na koži podeljeno sa trenutnom strujom koja prolazi kroz kožu (I):

$$R = \frac{V}{I}$$

GSR senzor sprovodi konstantnu voltažu kroz kožu preko elektroda. Voltaža je jako mala, neškodljiva i osoba koja učestvuje u eksperimentu ne oseća nikakve posledice. Izlazna vrednost je provodnost izražena u mikrosimensima. Elektronska šema GSR senzora je jednostavna (Slika 17) i sastoji se od jednog otpornika, kondenzatora i EKG elektroda.

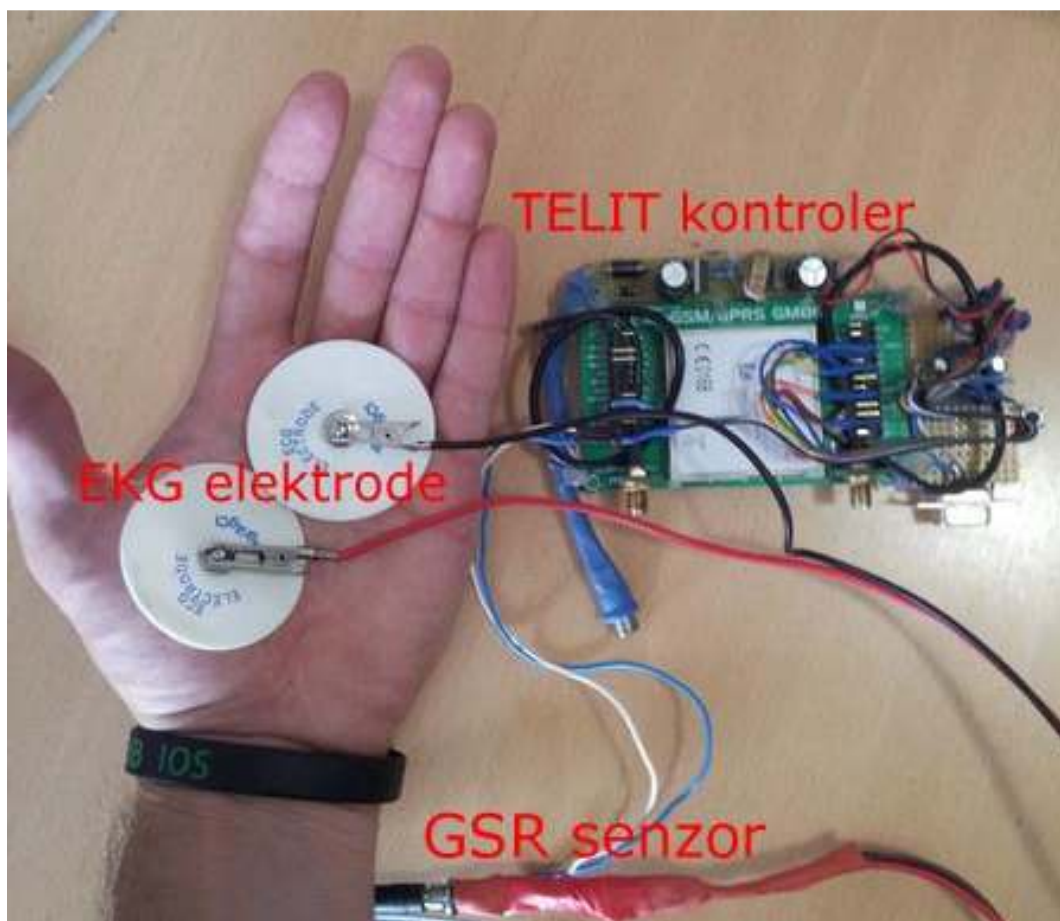


Slika 17 Šema GSR senzora

Samolepljive elektrode su korišćene da bi se obezbedila komforna upotreba senzora i da bi merenje otpornosti kože bilo konstantno i sa većom tačnošću.

Kod ljudskog tela, tkivo sa najvećom otpornošću je masno tkivo (nervi i mišići), dok je sama koža najbolji provodnik, zanemarujući mrtve ćelije epiderma (gornji sloj kože) koji je veoma loš provodnik. Zavisno od osobe do osobe, otpornost suve kože varira u vrednosti između 1,000-100,000 Oma. Najveću provodnost koža ima kada je mokra, što se i koristi kao metoda za određivanje provodnosti na koju direktno ili indirektno deluju psihološki uticaji tako što se povećava provodnost radi pojave fenomena znojenja (Fish & Geddes, 2003). EKG elektrode koje su korišćenje u eksperimentu imaju ulogu da održavaju konstantan kontakt sa dlanom osobe i provode struju.

Vrednosti koje zabeleži senzor prikupljene su preko TELIT-ovog GSM/GPRS kontrolera koji ima mogućnost logovanja vrednost na računar preko serial-to-usb interfejsa. Opcije slanja podataka preko GSM mreže koju ovaj uređaj nudi u ovom slučaju nije korišćena.



Slika 18 GSR senzor i kontroler za povezivanje na računar

Metoda upoređivanja studentovog ocenjivanja predavanja zasniva se ne upoređivanju provodnosti kože zabeležene u momentu kada je student ocenio predavanje kao interesantno ili neinteresantno. Upoređene su još i naknadne situacije u kojima je student učestvovao u interakciji.

3.3. Analiza i statistička obrada podataka

Značajnost razlika između aritmetičkih sredina ekstraktovanih iz interesantnih i neinteresantnih deonica testirana je pomoću t-testa za nezavisne uzorke gde nulta hipoteza glasi da nema razlike između istih parametara u interesantnim i neinteresantnim segmentima.

Međusobna zavisnost parametara izračunata je Personovim koeficijentom korelacije, koji pokazuje jačinu korelacije između varijabli kroz visinu korelacionog koeficijenta. Pomoću ove metode moguće je utvrditi značaj pozitivne ili negativne korelacije između parametara ekstraktovanih iz zvuka i pokreta u auditorijumu.

Tabela 9 Značenja vrednosti korelacionih koeficijenata

Korelacioni koeficijent		Značenje vrednosti
od	do	
.00	.19	Mala, gotovo zanemarljiva korelacija
.20	.39	Mala korelacija
.40	.69	Srednja korelacija
.70	.89	Visoka korelacija
.90	1.00	Veoma visoka korelacija

3.4. Metodologija evaluacije performansi sistema

Kao što je navedeno, glavni kriterijumi za evaluacije klasifikatora u ovom istraživanju su tačnost klasifikatora, broj tačnih pozitivnih klasifikacija (eng. True Positive), broj netačnih pozitivnih klasifikacija (eng. False Positive), preciznost (Witten et al. 1999) i kapa koeficijent.

Klasifikator predviđa kojoj klasi set podataka pripada sa četiri moguća ishoda: tačno pozitivan, tačno negativan, netačno pozitivan i netačno negativan. Korišćenjem ovih ishoda senzitivnost i specifičnost verovatnoće se konstruiše da bi se evaluirala sposobnost sistema da razlikuje pozitivne rezultate od negativnih.

Ako je tp definisan kao broj tačnih pozitivnih klasifikacija neuronske mreže, fn broj netačnih negativnih klasifikacija, fp broj netačno pozitivnih klasifikacija i tn tačno negativnih, senzibilnost je data kao:

$$r = \frac{tp}{tp + fn} 100\%$$

Specifičnost je definisana sledećom formulom:

$$q = \frac{tn}{tn + fp} 100\%$$

Postitivne vrednosti predikcije računaju se kao:

$$p = \frac{tp}{tp + fp} 100\%$$

Ukupna tačnost računa se na sledeći način:

$$t = \frac{tp + tn}{tp + tn + fp + fn} 100\%$$

Kapa koeficijent (Galron, 1982) je mera klasifikovanih N itema u C zajednički ekskluzivnih kategorija. Kapa se računa na sledeći način:

$$k = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

Gde je $\Pr(a)$ relativno mera poklapanja između ocenjivača i $\Pr(e)$ je hipotetička verovatnoća poklapanja, kada se posmatrani uzorak koristi za računanje verovatnoće za svaki posmatrani period nasumično za svaki kategoriju.

4. Zahtevi sistema za praćenja aktivnosti studenata

Kako bi se uspešno izašlo u susret problematici prepoznavanja aktivnosti studenata, neophodno je da se ispune određeni zahtevi i sagledaju ograničenja koja utiču na proces prikupljanja, analize i prezentacije signala pomoću IoT-a. Omogućavanje inteligentnih servisa u okruženju kakvo je učionica zahteva da se obezbede osnovne informacije o prostoriji (temperatura, svetlost, vlažnost vazduha), auditorijumu (broj učesnika) i interakciji između ljudi (pokreti, orijentacija tela i glave, reči koje izgovore međusobno, itd). U ovoj disertaciji detaljno su obrađene informacije iz ambijenta, odnosno pokreti i zvučni signal.

Praćenje aktivnosti studenata je grupno merenje koje zahteva istraživanje metode koja bi dala efikasne rezultate u približno realnom vremenu. Naglasak je na “približno” zbog jednostavne potrebe da prođe određena deonica predavanja koja se analizira, na primer deonica od 30 sekundi ili 1 minuta. Na ovom vreme neophodno je dodati i vreme obrade signala koji bi trebalo da bude zanemarljivo.

Prikupljeni podaci zahtevaju klasifikaciju, opis i skladištenje prema izvoru (parametri iz slike, zvuka, i sl). Merenje socioloških signala u grupi se može uprostiti praćenjem parametara koji su karakteristični za određeno sociološko ponašanje, ili direktnom analizom parametara, njihovih izmerenih vrednosti i odnosa između vrednosti za konkretan model. U daljem tekstu detaljnije su opisani zahtevi sistema kao i ostali zahtevi koje je neophodno ispuniti.

4.1. Sistemski zahtevi

4.1.1. Razlike između trenažnog i izvršnog okruženja

Uslovi u kojima se vrši prikupljanje podataka potrebno je da bude što sličniji okruženju u kome će sistem vršiti praćenje aktivnosti studenata. Pod sličnim podrazumeva se sličnost ambijenta učionice, mesta postavljanja IoT uređaja; a poželjno je i da broj studenata bude približan broju koji je korišćen za treniranje sistema. Ovi faktori mogu da se odraze na tačnost, ali nisu i presudan faktor. Određeni parametri koji imaju jaku statističku značajnost za određivanje aktivnosti studenata (npr. visok nivo buke i/ili pokreta), mogu se upotrebiti kao direktan indikator čime će se pojednostaviti prepoznavanje konteksta i greške svele na minimum.

4.1.2. Performanse sistema

U ovoj studiji performanse su evaluirane simulacijom prema različitim ulaznim parametrima, pre svega broja studenata i različitih predavača. Zahtev da se u realnom vremenu uradi klasifikacija je nemoguć iz prostog razloga jer je potrebno da prođe minimalan period predavanja za koji će se izvršiti snimanje deonica, segmentacija zvuka, razvrstavanje merenja i klasifikacija. Prema tome, ako se vrši prepoznavanje paterna za vreme predavanja, potrebno je zabeležiti deonicu predavanja, npr. proteklih 30 sekundi i za istu vršiti proces prepoznavanja. Kako bi se uradila prethodna segmentacija deonica zvuka na delove sa govorom i bez govora, potrebni su izlišni procesorski i memorijski resursi da bi se smanjilo kompletno vreme segmentacije i klasifikacije tako da prođe što manje vremena od započetog procesa rada sistema do narednih rezultata.

4.1.3. Simultano izvršavanje radnih procesa

Kao bitan faktor pri procesiranju signala predstavlja paralelno izvršavanje procesa koje se nameće zbog potrebe simultanog procesiranja slike, zvuka, i drugih merenja koja potiču od različitih senzora. Neophodno je u datom momentu uraditi analizu različitih signala a da bi se to omogućilo, za određene deonice predavanja u istom momentu porebno je uraditi procesiranje različitih zadataka (slike, zvuka, itd) da bi se sam proces analize deonica nesmetano mogao odvijati repetativno za vreme trajanja kompletnog predavanja.

4.1.4. Hardverski sistemski zahtevi

Proces obrade informacija prikupljenih iz ambijenta na prvom mestu zavisi od postavke sistema i načina na koji će se prikupiti i obraditi digitalan signal. Ukoliko se snimanje slike vrši u intervalu od jedne sekunde, hardver mora da podržava Burst Mode, odnosno snimanje slika u zadatom intervalu. Običan aparat ili kamera nemaju fiksni režim snimanja, što može pri pravljenju frejmova direktno sa kamera da izazove kašnjenja koja nisu fiksna zbog čega se ne može raditi snimanje u zadatom intervalu. Kvalitet snimljene slike i zvuka neophodno je da bude što bolji ali i da postoji što manje razlika u kvalitetu i prediscpozicijama koje su korišćene u setu za klasifikaciju neuralnih mreža i setu koji se koristi za prepoznavanje parametara, odnosno u izvršnom okruženju sistema.

4.2. Integritet sociološke interakcije

4.2.1. Mesto postavljanja IoT uređaja

Složeni sociološki procesi koje je neophodno zabeležiti zahtevaju posebne uslove, veliki broj senzora ili narušavaju prirodnost interakcije. Veliki broj upadljivih uređaja postavljenih u neposrednoj blizini studenta prete da utiču na integritet sociološkog

ponašanja u učionici, zbog čega je neophodno koristiti tehnike koje manje narušavaju ambijent snimanjem celokupne scene pomoću jedne kamere i izbegavanjem postavljanja senzora direktno na studente. Psihološki efekat usled posmatranja može da bude takav da značajno utiče na ponašanje studenata što se može odraziti na tačnost prepoznavanja parametara iz ambijenta.

Kada su studenti svesni da su posmatrani, njihovo ponašanje neće biti uobičajeno, čak je moguće da se negativno utiče na njihovu pažnju, jel će njihova pažnja biti fokusirana na uređaje postavljene u njihovoj neposrednoj blizini. Delemično se može izbeći ovaj psihološki efekat ako se IoT uređaji za snimanje postave tako da ih studenti ne vide. Međutim samo prisustvo pametne prostorije u edukativnoj ustanovi i svest o njenoj nameni uvek će biti prisutna i jednim manjim procentom uticaće na ishod merenja. Pretpostavka je da će se ovaj procenat smanjivati vremenom, kako učesnicima ambijent bude postajao uobičajen.

4.3. Skladištenje i protok informacija

Proces prikupljanja podataka je praćen potrebom da se izvrši skladištenje određenog broja informacija na jednom mestu. Obzirom da izlazna vrednost sistema može da sadrži samo jedan parametar koji se odnosi na vrednost kvaliteta nastave date kao zainteresovanosti studenata. Kvalitativno projektovanje sistema može da smanji broj prenesenih, kao i broj smeštenih informacija ukoliko se koriste efikasniji formati za prenos (Gligorić et al. 2011). Ova problematika nije toliko izražena kod servera, koliko je to bitno kod senzora koji imaju ograničene resurse po pitanju memorije, baterije i procesorske snage.

4.4. Otvorena arhitektura

Prikupljeni paterni sadrže vredne informacije koje se mogu dalje upotrebljavati od strane različitih platformi ili koristiti u svrhu novih istraživanja u društvenim i računarskim naukama. Prema tome, određena otvorenost pristupa podacima bez potrebe

dodatnih podešavanja sistema i intervencije krajnjeg korisnika uz podrazumevanu autorizaciju je ključna kada se čitanje podataka vrši preko različitih kanala. Model sistema koji bi omogućio ovakve usluge baziran je najčešće na Cloud computing tehnologiji koja dozvoljava IT usluge preko Internet protokola, i obično uključuje dinamičke, skalabilne i virtualizovane resurse.

4.5. Presentacija podatka

Proces analize podataka i prezentacije je potrebno obaviti u realnom vremenu. Informacije prikupljne pomoću IoT-a neophodno je prezentovati krajnjem korisniku i potencijalnom skupu novih korisnika. Krajnji korisnik (predavač) koji može da prati aktivnost upotrebom pametnog telefona ili preko računara pomoću aplikacije za mobilni telefon ili nekog online servisa za vizualizaciju kao što su Cosm (Cosm, 2014), Google Charts (Google Chart Tool, 2014) ili Microsoft Pilot Viewer (PilotViewer, 2014). Cosm omogućava pored vizuelnog prikaza podataka i povezivanje novih korisnika na sistem (druge platforme ili aplikacije) posredstvom otvorenog interfejsa. Od momenta kada podaci merenja budu prisutni na Internetu, broj načina na koji se oni dalje mogu upotrebiti je neograničen.

5. Obzervaciono istraživanje i prikupljanje podataka za klasifikaciju

Kako bi se sagledala mogućnost praćenja aktivnosti na osnovu obrade statistički relevantnih parametara, tako da se za dalji tok istraživanja i donošenja finalnih zaključaka uopšte, koriste primenjive naučne metode, urađeno je obzervaciono istraživanje i anketiranje studenata. Na kraju je napravljen presek sa potencijalnim parametrima i izveden finalan skup parametara, upotrebljene su metode anketiranja i posmatranja studenata. Anketiranje je za rezultat dalo sa stanovišta studenata uvid u uticaje na kvalitet predavanja, kao i način ispoljavanja nezainteresovanosti.

5.1. Anketiranje uzorka

Na uzorku od 300 studenata sa Fakulteta organizacionih nauka i Saobraćajnog fakulteta, Univerziteta u Beogradu, sprovedena je anketa s ciljem da se utvrdi ponašanje studenata za vreme predavanja kada im predavanje nije interesantno. Izuzev ponašanja, pitanja su bila tako formirana da se odrede i spoljašnji faktori koji u određenoj meri mogu da utiču na ishod predavanja.

5.1.1. Rezultati ankete

Označite šta po Vama ima najveći uticaj na kvalitet predavanja:	veliki uticaj	ima uticaj	ne znam	mali uticaj	nema uticaj
Spoljasnja buka	16.5	50.4	1.3	27.0	4.8
Monoton glas predavača	60.0	29.1	1.7	6.5	2.6
Prejaka ili preslaba osvetljenost	15.7	54.8	4.3	19.1	6.1
Neprikladna temperatura	24.3	47.4	2.2	21.7	4.3
Veliki broj studenata u učionici	30.9	40.0	8.7	14.3	6.1
Koliko bi trebalo da traje optimalan čas	90 min	60 min	45 min	30 min	ne znam
	2.6	5.7	64.3	25.7	1.7
Deo dana kada najlakše pratite predavanja	jutro	podne	poslepodne	veče	ne znam
	50.9	41.7	5.7	0.9	0.9
Dan u nedelji najlošiji za predavanja	ponedeljak	petak	od utorka do vikend	ne znam	
	9.6	19.6	8.7	58.3	3.9
U kom godišnjem dobu najlakše pratite predavanja	proleće	leto	jesen	zima	ne znam
	32.2	3.0	22.2	28.3	14.3
U kom godišnjem dobu najteže pratite predavanja?					
	12.6	49.1	10.0	18.3	10.0

Slika 19 Rezultati ankete 2-1

Rezultati ankete pokazuju da su generalno pokreti i zvuk dva kanala kroz koje studenti najčešće ispoljavaju svoje nezadovoljstvo i nezainteresovanost za vreme predavanja. Studenti reaguju na monotona predavanja sa određenim paternima u ponašanju: upuštaju se u komunikaciju sa kolegama, pomeraju se na stolici, crtaju po papiru, koriste mobilni telefon, igraju se sa raznim predmetima (gumicom, olovkom), zevaju; tragaju za vizualnim stimulansima gledajući kroz prozor, itd. Sve ove akcije proizvode paterne koji se mogu dovesti u vezu sa pažnjom studenata i njihovom zainteresovanošću sa predavanjem.

Kada Vam je predavanje dosadno, kako to ispoljavate?	uvek	često	povremeno	nikad	ne znam
Pričam sa kolegama	5.2	30.0	57.8	5.2	1.7
Gledam kroz prozor	1.7	20.9	53.9	21.7	1.7
Vrpoljim se na stolici	13.0	27.8	39.6	17.4	2.2
Crtam po papiru	17.0	25.2	38.7	18.3	0.9
Koristim telefon	6.1	29.6	45.2	19.1	0.0
Igram se sa predmetima (gumicom, olovkom...)	10.4	19.6	38.3	29.1	2.6
Pravim se da pratim predavanje iako mi je dosadno	8.3	24.3	49.6	16.1	1.7
Zevam	7.4	39.6	46.1	4.3	2.6
Surfujem internetom	3.9	20.0	29.1	43.0	3.9

Slika 20 Rezultati anekete 2-2

Interesantno je istaknuti da je veliki broj studenata naveo da buka, osvetljenje, temperatura, broj učesnika predavanja u učionici; godišnje doba i deo dana imaju veliki uticaj na predavanje. Ovi parametri prevazilaze obim ove disertacije i neće biti obrađivani, ali svakako predstavljaju jednu od mogućih tema za dalje istraživanje.

5.2. Posmatranje uzorka i izgradnja dataset-a

Naučna metoda posmatranja sprovedena je na uzorku od 15 studenata koji su snimani kamerom za vreme predavanja. Ukupan broj seansi bio je 20 školskih časova (15h), od kojih je svaki dužine 45 minuta. Kamera je instalirana u uglu učionice, snimajući samo studente (Slika 1). Zvučni signal je takođe sniman kamerom, čime je zabeležen i ambijetalan zvuk zajedno sa ehoom u učionici. Cilj je bio da se snimi prirodna mešavina zvukova, ali bez spoljnih faktora koji se ne mogu kontrolisati (npr. spoljašnja buka), tako da trenažno okruženje bude što sličnije okruženju u kome će sistem da izvršava procese prepoznavanja parametara. Algoritam koji je korišćen za detektovanje

parametara (Basu, 2002) iz signala je robustan na distancu postavljanja mikrofona i na potencijalnu buku koja se može pojaviti, čime je moguće sprovesti istraživanje u pomenutim uslovima.



Slika 21 Scena zabeležena kamerom za vreme istraživanja

Sistem takođe obuhvata i praćenje pokreta predavača i korelaciju sa ostalim parametrima, ali je ovaj deo istraživanja u sklopu svoje doktorske disertacije obradila kolegica Ana Uzelac.

5.2.1. Anotacija podataka

Snimana predavanja su bila interaktivnog tipa, što znači da su studenti aktivno učestvovali u predavanju, odgovaranjem na pitanja predavača, kako bi se održala zainteresovanost auditorijuma na visokom nivou.

Anotacija podataka je proces povezivanja događaja sa korespondiranim aktivnostima.

U ovom istraživanju anotacija podataka je urađena kako bi se omogućilo treniranje mašinskog algoritma korišćenjem različitih klasa podataka. Segmenti od 30 sekundi su označeni kao jedna od dve klase: “interesantni“ ili “neinteresantni“. Proces anotacije se može uraditi manuelno od strane istraživača – metoda koja je generalno manje tačna i

podložna greškama, ili od strane studenata. Takođe, anotacija se često obavlja tako što se od korisnika, u ovom slučaju je to student, zamoli da replicira određenu aktivnost.

Za vreme nastave studenti su koristili mobilnu aplikaciju za ocenjivanje predavanja da bi izrazili svoju zainteresovanost (nezainteresovanost) za određene delove prezentacije upotrebom “Like” i “Dislike” dugmeta. Studentima su date instrukcije da ocene predavanje predavača svaki put kada im je nešto posebno interesantno ili neinteresantno. Kao podsetnik, aplikacija za ocenjivanje na mobilnim uređajima je izbacivala poruku svakih 60 sekundi, podsećajući studente da glasaju.



User number: 15
[idi na anketu](#)

Slika 22 Aplikacija za ocenjivanje predavanja

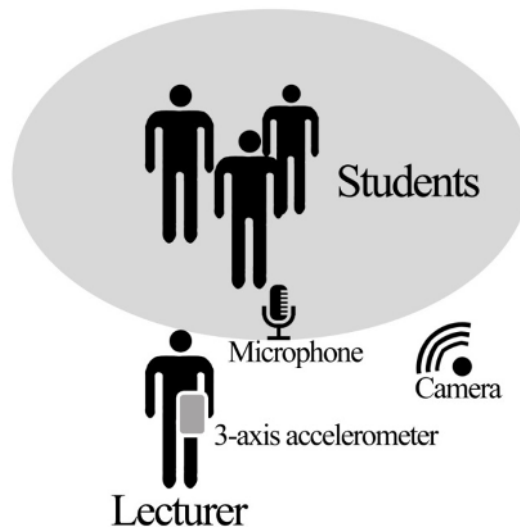
Likertova skala je pouzdaniji način da se odredi stepen zainteresovanosti studenata pri glasanju, međutim povećan broj opcija uticao bi na sposobnost studenata da prate nastavu i u istom vremenu ocenjuju predavanje. Prema tome, aplikacija se sastoji samo od dva tastera *like* i *dislike*.

Student na koga je bio prikazan GSR uređaj, imao je zadatak što manje da se pomera da bi GSR očitavanja bila što tačnija zbog položaja elektroda na ruci.

Ukupno četiri grupe studenata su učestvovala u ocenjivanju predavanja i za to vreme kamerom i mikrofonom snimana su predavanja u učionici. Broj studenata u svakoj grupi bio je sledeći:

- Grupa 1 – 11 studenata
- Grupa 2 – 12 studenata
- Grupa 3 – 13 studenata
- Grupa 4 – 10 studenata

Kamera je postavljena u ugao učionice snimajući samo studente, dok je širokopojasni Rode NT2A mikrofonski postavljena na centar učionice. Android telefon sa ugrađenim 3-axis akcelerometrom je prikačen na predavača. Zvučni zapis snimljen pomoću mikrofona uključuje i ambijentalan zvuk pored glasova studenata i predavača. Cilj je bio da se zabeleži prirodna mešavina zvučnih izvora tako da trenajni dataset bude što sličniji izvršnom koliko je moguće.



Slika 23 Postavke senzora i opreme pri obzervacionom istraživanju

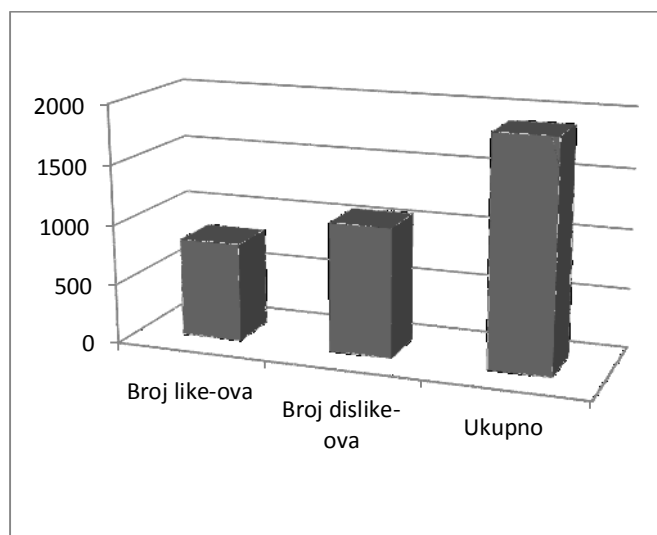
Ukupan snimljeni materijal od 15 sati je segmentiran na sledeći način zavisno od glasova u svakoj deonici od 30 sekundi:

- (1) deonice sa svim negativnih ocenama su označene kao “neinteresantne”
- (2) deonice sa svim pozitivnim ocenama su označene kao “interesantne”
- (3) više od jednog glasa od strane jednog studenta u jednoj deonici su zanemareni
- (4) deonice sa malim procentom ocena studenata (<90%) su zanemarene

Proces anotacije podataka je sproveden na sledeći način:

- Vremenska razlika vremena na serveru i vremena kada je fajl snimljen je sinhronizovano oflajn
- Slike u kojima je predavač slučajno ušao u kadar su izbačene iz seta
- Druge sekvence sa nepredviđenim aktivnostima, npr sa bukom koja dolazi van učionice; kada studenti ulaze u učionicu za vreme predavanja su takođe izbačene (<1)

Nakom procesa prethodno opisanih, dataset je sveden na 60 minuta sa deonicama dužine trajanja od 30 sekundi označenih kao “interesantne” i “neinteresantne”. Dužin trajanje deonice je izabrana na osnovu pilot istraživanja koje je sprovedeno i u kom je eksperimentisano sa dužinom trajanja deonice. Algoritam je pokazao najveću tačnost kada su deonice bile dužine 30 sekundi.



Slika 24 Prosečan broj glasova po deonici i ukupan broj glasova u datasetu

Prosečan broj glasova po sekvenci za grupe 1,2,3 i 4 bio je 10.4, 11.2, 12.1 i 9. Predloženi metod za anotaciju podataka zahteva veliki broj glasova da bi se izgradio dataset zbog malog broja glasova koji studenti generišu. U ovom istraživanju od 15 sati

snimljenog materijala napravljeno je 1h dataseta. Druga moguća metoda za prikupljanje podataka bila bi da se od studenata traži da ocene predavanja u određenom momentu čime bi se omogućilo prikupljanje većeg broja glasova u svakoj deonici, ali na ovaj način bi se poremetila prirodnost sociološke interakcije u učionici i čime obzervaciono istraživanje može dati pogrešne rezultate. Na slici iznad je prikazan broj glasova u datasetu za određene grupe studenata kao i ukupan broj glasova za dve klase i ukupan broj glasova u celom datasetu. Broj glasova u označenim sekvencama je prikazan u tabeli i ovako sortiran predstavlja dataset koji će biti korišćen za evaluaciju klasifikatora.

Tabela 10 Karakteristike dataseta nakon segmentacije

	Grupa 1	Grupa 2	Grupa 3	Grupa 4	Total
Broj deonica	20	30	40	30	120
Broj studenata	11	12	13	10	46
Prosečan broj glasova po deonici	10.4	11.2	12.1	9	10.7
Ukupan broj glasova po deonici	208	224	484	270	1186
Broj deonica označenih kao interesantne	15	21	26	17	79
Broj interval označenih kao neinteresantne	5	9	14	13	41

5.2.3. Rezultujući set podataka

Rezultat obzervacionog istraživanja je 15 sati video materijala, što je posle segmentacije (samo deonice sa velikim procentom like ili dislike ocena su uzete u obzir) svedeno na 1h sat uzorka sa obeleženim interesantnim i neinteresantnim deonicama. Deo materijala je odbačen pre segmentacije iz razloga što je jedan od zahteva sistema da se utvrdi eksperimentalna kontrola u okruženju kakva je učionica (npr. diskusije koje nisu u vezi sa temom predavanja su odbačene; često se dešavalo da neko napusti ili uđe u učionicu, tako da su ovi delovi uklonjeni iz seta).

Iz segmentiranog materijala, dalje su ekstraktovani parametri iz zvuka i slike. Iz zvuka su potom ekstraktovani i analizirani parametri koji se obično koriste kod analize vokalnog ponašanja: glasovni segmenti, govorna aktivnost, formantna frekvencija, spektralna entropija, autokorelacija i stopa glasova (ili glasovna stopa).

Pokreti studenata ekstraktovani su sa video frejmova merenjem strukturalnih razlika između dva naredna frejma u svakoj sekundi upotrebom DSSIM algoritma robustnog na promene osvetljenja u ambijentu.

Galvanska provodnost kože, kao mera spoljašnjih uticaja na pojedinca merena je kod jednog studenta radi utvrđivanja verodostojnosti ocene predavanja koju je student dao za vreme posmatranja studenata. Provodnost kože je merena manuelno napravljenim GSR uređajem, koji prilikom puštanja slabe struje kroz kožu meri promene otpornosti, što je direktan pokazatelj promena u provodnosti kože i prema tome pokazatelj realnog uticaja predavanja na pojedinca.

6. Analiza parametara

Za svaki parametar ekstraktovane su srednje vrednosti, standardna devijacija i izračunata je statistička značajnost preko T-Testa za nezavisne uzorke. Od osnovnih izvedenih parametara iz zvuka, modelovani su složeni sociološki signali konverzacioni aktivnosti i stres. Potom da bi se odradila korelacija između svih karakteristika za interesantne i neinteresantne deonice, izračunat je Pearsonov koeficijent korelacije.

6.1. Parametri iz zvučnog signala

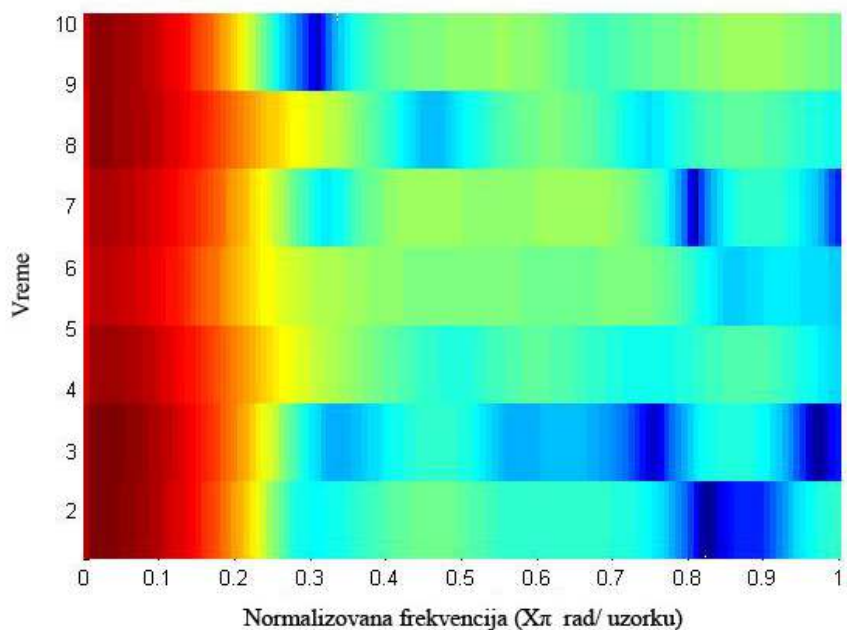
Tabela 11 Parametri ekstraktovani iz zvuka

	Srednja vrednost		Standardna devijacija		T-Test	Značajnost
	Like	Dislike	Like	Dislike		
<i>Glasovni segmenti</i>	2.84	1.94	1.34	0.719719159	2.58	p = 0.015
<i>Govorna aktivnost</i>	0.025	0.0011	0.0011	0.00043	4.64	p = 9.93 ⁻ 07
<i>Formantna frekvencija</i>	216.62	237.99	23.59	18.53	-3.11	p = 0.0038
<i>Spektralna entropija</i>	3.75	3.7	0.079	0.048	2.28	p = 0.029
<i>Vrednost najvećeg autokorelacionog pika</i>	0.69	0.71	0.0412	0.0465	-1.83	p = 0.08
<i>Lokacija najvećeg autokorelacionog pika</i>	0.0241	0.0261	0.0022	0.002032398	-2.80	p = 0.008
<i>Broj autokorelacionih pikova</i>	7.00	7.28	0.395673453	0.326423746	-2.35	p = 0.025
<i>Stopa glasova</i>	0.027	0.031	0.0063	0.0054	-1.8	p = 0.08

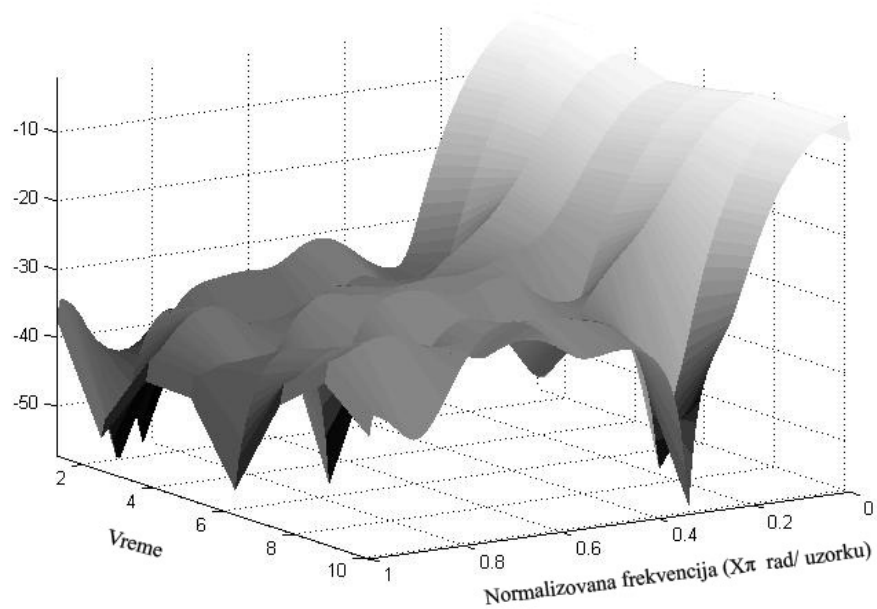
Glasovni segmenti su dobar pokazatelj zainteresovanosti, što se može zaključiti na osnovu male p vredosti ($p = 0.015$) gdesu za svega 1.5% celokupnog skupa podaci bili slični, kao i velike razlike između deonica gde je ekstraktovano 68% više govornih segmenata za delove predavanja označene interesantnim.

Govorna aktivnost je još bolji statistički pokazatelj pažnje ($p = 9.93^{-7}$), sa gotovo dvostruko većim vrednostima govorne aktivnosti za interesante delove predavanja. Srednje vrednosti za govornu aktivnost su dosta male iz razloga što je detekcija govora iz zvučnog zapisa na kome se nalazi veći broj ljudi otežana; broj identifikovanih deonica kao govor je manje što je više preplitanja govora, što je pokazatelj opadanja pažnje auditorijuma.

Formatna frekvencija - Spektrogram formantne frekvencije je izveden pomoću brze Furijeove transformacije za sve sekvence kako bi se analizirale vokalne informacije iz zvuka u učionici i pokazale razlike između formatne frekvencije za interesantne (slike 25 i 26) i neinteresantne deonice (slike 27 i 28). Statistički posmatrano ($p = 0.0038$), znatno veća formatna frekvencija (91%) zabeležena je u auditorijumu kada studenti ne prate predavanje, odnosno kada im predavanje nije interesantno. Svetlija boja na slikama označava više akustične energije u govornim segmentima.

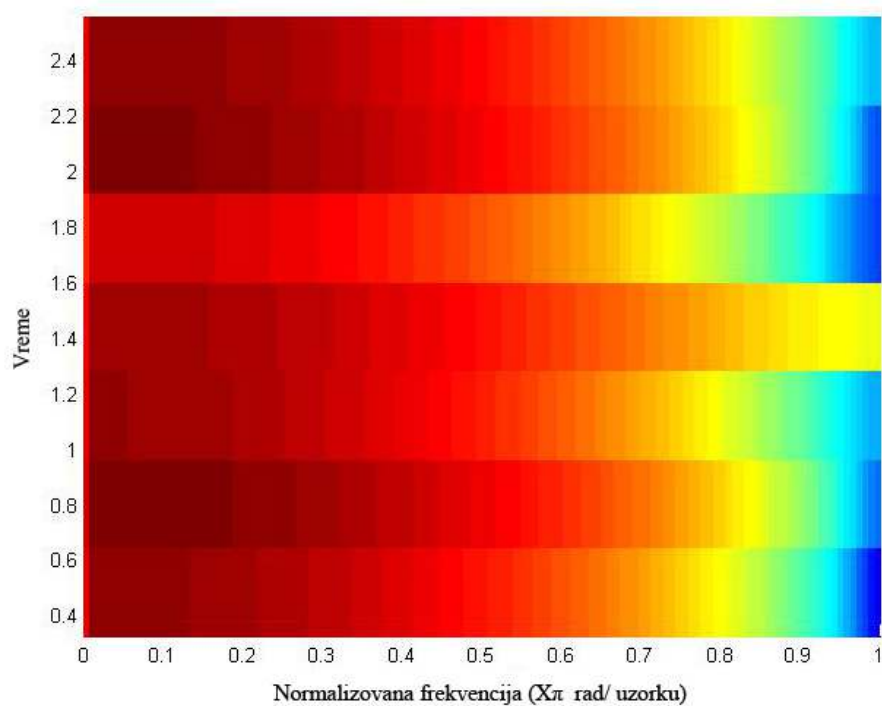


Slika 25 Spektrogram normalizovane frekvencije u vremenu za interesantne deonice

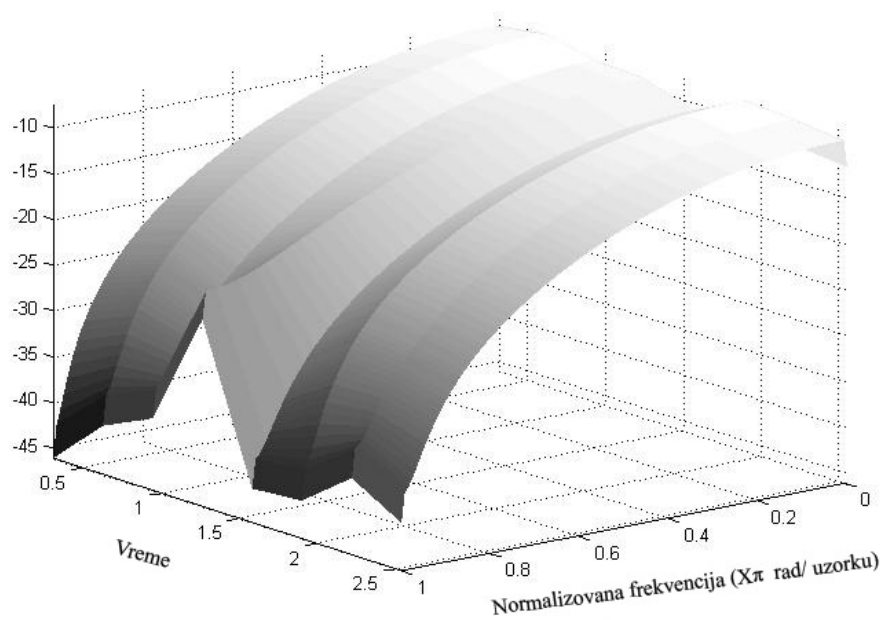


Slika 26 Spektrogram formatne frekvencije za interesantne sekvence

Povećan nivo energije u frejmu je često znak dominantnosti, ali u slučaju kada se snimanje zvuka obavlja pojedinačnim mikrofonom postavljenim na malom odstojanju od glave govornika. U ovom slučaju, kada postoji jedan signal na kome se nalaze svi učesnici u interakciji, drugačije se posmatraju i tumače rezultati, jel je veliki broj vokala detektovan u kratkom periodu i sama energije koja se manifestuje tom prilikom, kako se može primetiti sa spektrograma – pokazatelj nezainteresovanosti.



Slika 27 Spektrogram normalizovane frekvencije u vremenu za interesantne deonice



Slika 28 Spektrogram formatne frekvencije za neinteresantne sekvence

Autokorelacija (najveća vrednost, pozicija i broj autokorelacionih pikova) koristi se za modelovanje buke, visine tona i eha iz zvučnog signala. Ovaj parametar je važan jer se koristi pri određivanju segmenata u kome ima govora. T-test je pokazao da autokorelacioni parametri, izuzev broja autokorelacionih pikova u svakoj sekvenci ($p = 0.025$), nisu najbolji indikatori zainteresovanosti studenata. Međutim sam broj pikova direktno je pokazatelj da je u zabeleženom signalu prisutan govor. Sa druge strane, autokorelacija i gore navedeni parametri zajedno formiraju različite sociološke signale, pogotovo zahvaljujući spektralnoj entropiji koja sama po sebi nije od značaja, ali zajedno sa formantnom frekvencijom omogućava detekciju složenih procesa. U slučaju kada postoji više visokih tonova, što je mereno pozicijom i brojem autokorelacionih pikova, veće su šanse da studenti ne prate predavanje. U suprotnom povećanje visine tona u glasu može da dolazi od predavača koji podiže glas da bi privukao pažnju studenata, što je opet znak postojanja nezainteresovanosti.

Stopa glasova – predstavlja broj glasova u određenoj deonici za koji se u ovom istraživanju pokazalo da nije statistički relevantan parametar: T-test za glasovnu stopu nije pokazao statističku značajnost ($p = 0.08$).

Stres – interesantne i neinteresante sekvence su u negativnoj korelaciji: korelacija stresa je bila negativna i više nego dvostruko veća za delove predavanja koje su studenti ocenili kao neinteresantne. Kao što je ranije navedeno, stres može da bude nameran (govornik namerno naglašava deonice govora), ili nenameran (psihološki stres izazvan nelagodnošću), i takođe određuje dinamiku govora, čime više stresa može da označava dinamičniji govor. Imajući u vidu da su predavanja bila interaktivnog tipa, može se zaključiti da za vreme interaktivnih predavanja emituje više pozitivnog nenamernog stresa, sa željom da se intonacijom privuče pažnja studenata i da je govor bio dinamičniji. Osnovni parametri koji se upotrebljeni za modelovanje stresa su statistički značajni.

Tabela 12 Ukupan stres zabeležen kod interesantnih i neinteresantnih sekvenci

Sekvence	Ukupan stres
Interesantne	-0.15051
Neinteresantne	-0.40601

Konverzaciona aktivnost je bila 11.43% veća za interesantne deonice, ali se pokazalo da ovaj parametar nije statistički značajan ($p=0.2$).

Tabela 13 Ukupna konverzaciona aktivnost za interesantne i neinteresantne sekvence

Sekvence	Ukupna konverzaciona aktivnost
Interesantne	2.673632857
Neinteresantne	2.368042105

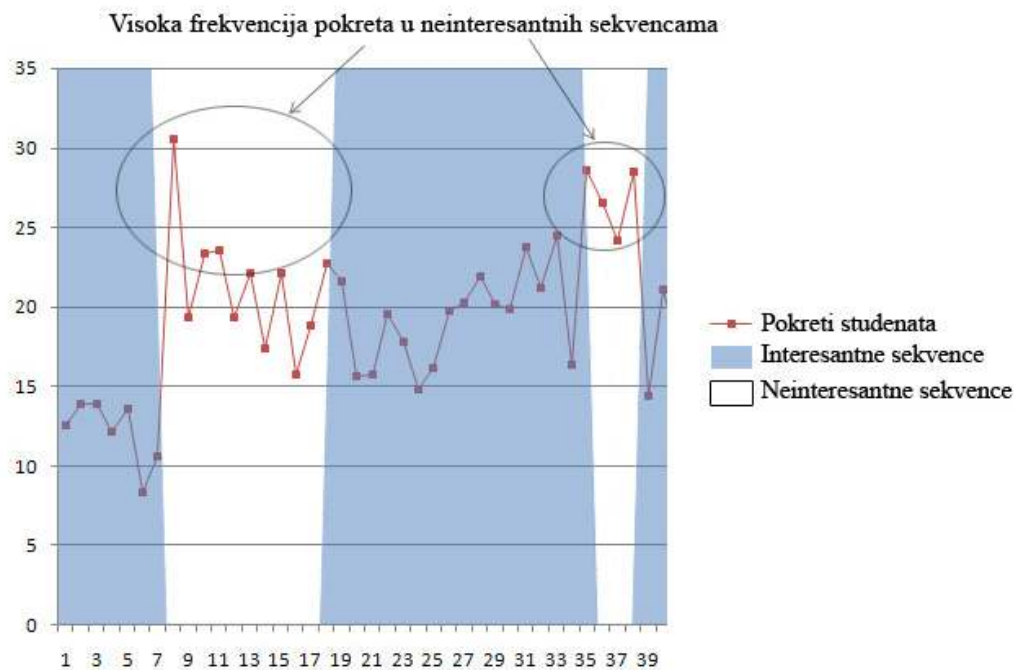
Prethodni rezultati opisani u ovom poglavlju potvrđuju hipotezu H3a, tako da se može zaključiti da se određeni parametri iz zvuka i pokreta mogu iskoristiti za identifikaciju pažnje studenata.

6.2. Parametri ekstraktovani sa slike

Tabela 14 Pokreti studenata ekstraktovani indeksom strukturalne razlike DSSIM između frejmova

Pokreti studenata	Interesting sequences	Not-interesting sequences
Srednja vrednost	16.33	21.68
Standardna devijacija	4.098	3.588
T-Test	t = -4.28	p= 14 ⁻⁵

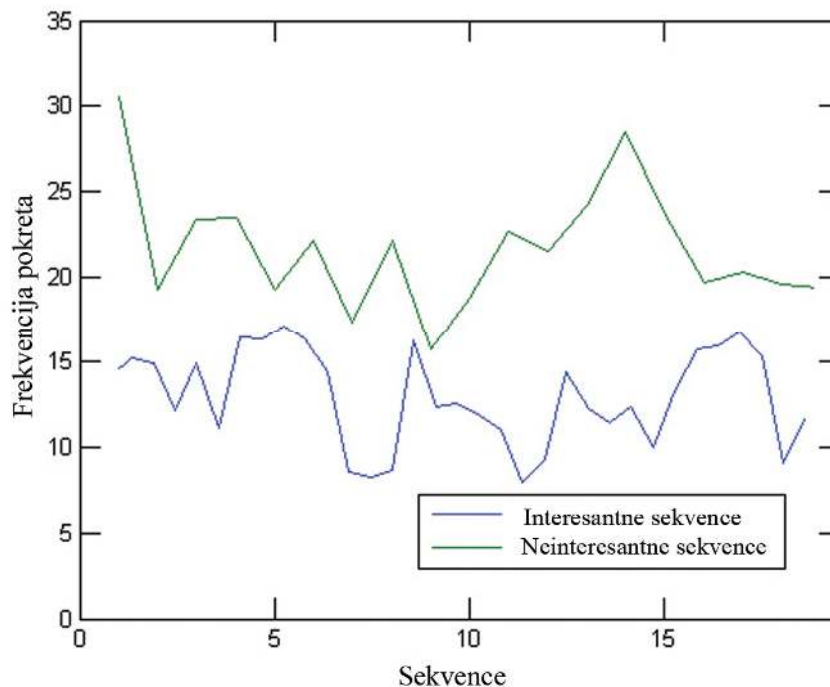
Srednja vrednost i standardna devijacija ekstraktovanih strukturalnih razlika iz frejmova koji predstavljaju frekvenciju pokreta studenata, pokazuju da su pokreti studenata češći i da se studenti više pomeraju kada im je predavanje dosadno – u ovom istraživanju 75% više.



Slika 29 Visoke vrednosti frekvencije pokreta za neinteresantne sekvence

T-statistika od -4.28 i značajno mala p vrednost (14⁻⁵), direktni su pokazatelji da je nezainteresovanost studenata u uskoj povezanosti sa povećanom frekvencijom pokreta u

auditorijumu i da studenti ispoljavaju nezadovoljstvo nesvesno se pomerajući znatno više nego kada im je predavanje interesantno.



Slika 30 Frekvencija pokreta u interesantnim i neinteresantnim deonicama

6.3. Galvanska provodnost kože

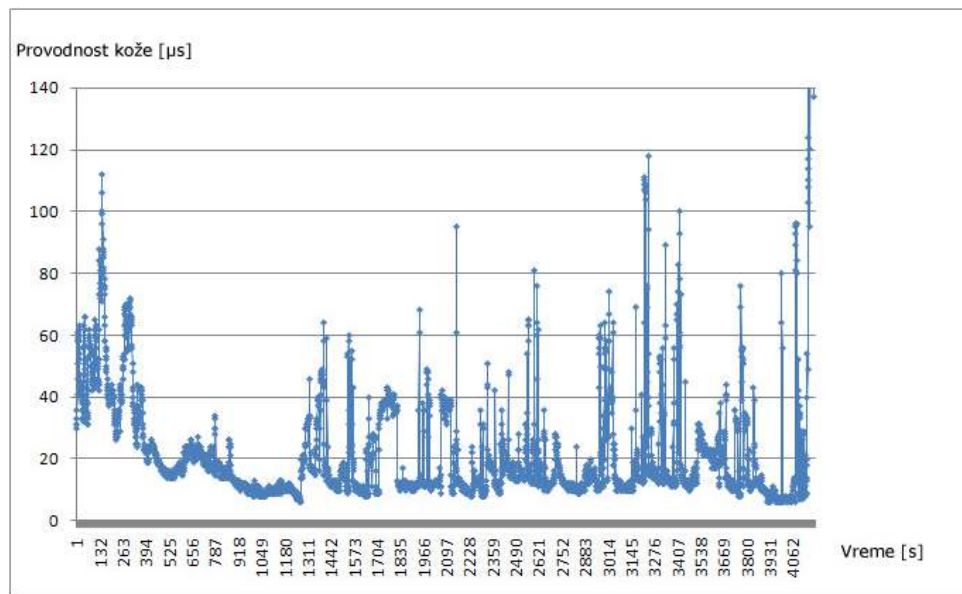
Provodnost kože nije pokazala značajne razlike za interesante i neinteresantne deonice. Nasuprot statističkim pokazateljima koji nisu bili merodavni za donošenje zaključka o zainteresovanosti studenata, merenja provodnost kože pokazala su vezu psihološkog uticaja okruženja na pojedinca. Kako bi se provodnost kože koristila za praćenje zainteresovanosti predavanja, neophodno je uključiti dodatne parametre, poput EKG signala. Cilj ovog istraživanja nije bio da se utvrdi veza između provodnosti kože i zainteresovanosti studenata; prema tome, dobijeni rezultati predstavljaju dodatnu vrednost. Zbog toga što nije statistički utvrđena značajnost za ovaj parametar, urađene su naknadne analize na malom uzorku GSR merenja sa napomenom da uzorak nije

relavantan ali predviđa interesantne rezultate koji se mogu koristiti kao smernica za buduća istraživanja.

Tabela 15 Galvanska provodnost kože u momentu ocenjivanja

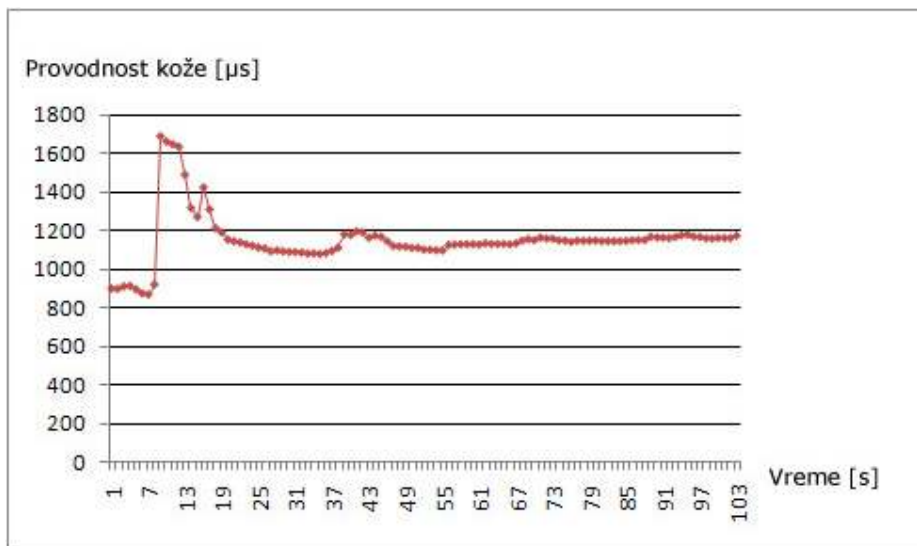
	<i>Srednja vrednost provodnosti kože</i>		<i>T-Test</i>	<i>Značajnost</i>
	<i>Like</i>	<i>Dislike</i>		
<i>Student 1</i>	130.17	124	0.88	0.38
<i>Student 2</i>	1168.33	1159.26	0.177	0.86

U slučaju kada student učestvuje u interakciji sa profesorom i drugim studentima, vrednosti koje je senzor zabeležio bile su dvostruko veće, što može biti znak da interaktivna predavanja bolje utiču na pažnju studenata, ali ova činjenica zahteva dodatna istraživanja, jer je uzorak bio mali.



Slika 31 Dijagram provodnosti kože za prvog studenta

Na slici ispod, sekvenca sa visokim vrednostima provodnosti kože prikupljena GSR senzorom prikazana je za period kada je student započeo interakciju sa predavačem. Ostale vrednosti su bile ravnomerno raspoređene bez značajnih statističkih razlika.



Slika 32 Dijagram provodnosti kože za drugog studenta

Visoke vrednosti za provodnost kože (Slika 32) daju jasan prikaz promene vrednosti kada se student upuštao u interakciju sa predavačem. Iz razloga jer su same promene provodnosti kože pokazatelj promene zainteresovanja, uzorci standardne devijacije za deonice od 30 sekundi testirani su korišćenjem t-testa za nezavisne uzroke da bi se utvrdilo da li su za interesantne i neinteresantne deonice različiti. Rezultati t-testa kao i vrednosti za deonice za standardne devijacije prikazane su u tabeli 14.

Tabela 16 Promene galvanske provodnosti kože kroz deonice od 30 sekundi i srednja standardna devijacija

	Srednja vrednost standardne devijacije		T-Test	Značajnost
	Like	Dislike		
Student 1	2.45	6.74	-1.67	0.106
Student 2	68.76	5.25	1.99	0.0586

Kod prvog studenta 10.6% uzoraka je bio sličan za različite klase deonica, dok je kod drugog ta sličnost iznosila oko 5.9% što za 0.9% prevazilazi granicu statističke značajnosti. Takođe, vrednosti koje su dobijene za standardnu devijaciju se razlikuju za interesantne i neinteresantne deonice za oba studenta: kod jednog studenta skoro tri puta veća devijacija je zabeležena kada student nije bio zainteresovan predavanjem, dok je standardna devijacija bila 13 puta veća kada je drugi student bio zainteresovan

predavanjem. Prema tome, dovođenje u vezu provodnosti kože sa ocenama koje su studenti dali za vreme predavanja pritiskanjem „like“ i „dislike“ dugmeta nije dalo kvalitativne rezultate. Ovo ne znači da ocene studenata nisu relevantne, već da se jednostavno GSR ne može koristiti kao metoda za određivanje pažnje zbog toga što promene provodnosti kože kada su studenti zainteresovani ili nezainteresovani nije samo po sebi dovoljno. U momentima kada je student imao interakciju sa predavačem provodnost kože je bila dvostruko veća, tako da je upravo interakcija način da se studentima pažnja održi na visokom nivou, što je moguća tema za dalja istraživanja.

6.4. Korelacija parametara

Tabela 17 Korelacija parametara za sve karakteristike: „neinteresantni“ parametri označeni su sivom bojom. *Parametri kod kojih je $p < 0.05$ za oba parametra čija je korelacija merena

	1	2	3	4	5	6	7	8	9	10	11
1. Glasovni segmenti	-	0.36	-0.94	0.01	0.43	0.35	-0.02	0.46	0.03	0.10	-0.07
2. Vreme govora	0.44	-	-0.51	-0.52	-0.21*	-0.15	0.19	-0.18*	0.02*	0.47	-0.57*
3. Glasovna stopa	-0.89	-0.53	-	0.10	-0.33	-0.34	-0.13	-0.33	-0.01	-0.16	0.23
4. Broj kratkih govornih segmenata	-0.28	-0.72	0.39	-	0.20	0.19	-0.25	0.17	0.01	-0.20	0.56
5. Formantna frekvencija	0.31	-0.27*	-0.28	0.18	-	0.66	0.15	0.94*	0.19*	0.20	0.49*
6. Pouzdanost formantne frekvencije	0.31	-0.04	-0.20	-0.03	0.51	-	0.64	0.62	0.41	-0.04	0.43
7. Vrednost najvećeg autokorelacionog pika	0.10	0.31	-0.07	-0.29	-0.16	0.58	-	0.10	0.15	0.08	0.21
8. Lokacija najvećeg autokorelacionog pika	0.33	-0.03*	-0.33	0.01	0.79*	0.62	0.23	-	0.30*	0.07	0.39*
9. Broj autokorelacionih pikova	-0.13	-0.35*	0.15	0.19	0.38*	0.23	-0.29	0.30*	-	0.05	-0.17*
10. Govorna aktivnost	0.13	0.43	-0.16	-0.19	-0.30*	-0.15	0.36	0.00*	-0.36*	-	-0.04*
11. Pokreti studenata	-0.10	0.20*	0.03	-0.28	-0.23*	0.02	0.26	-0.09*	-0.05*	-0.09	-

Vrednosti korelacija između različitih sekvenci su u opsegu [0, 0.94]. Za “interesantne” sekvence statistički značajna veća korelacija (p najmanje 0.05 za oba parametra i $r > 0.7$) zabeležena je između formantne frekvencije i pozicije najvećeg autokorelacionog pika ($r = .79$).

Za “neinteresantne” sekvence visoka korelacija je takođe bila slučaj kod formantne frekvencije i pozicije najvećeg autokorelacionog pika ($r = 0.94$). Srednja korelacija ($0.7 > r > 0.4$) je zabeležena kod vremena govora i pokreta studenata (-0.57); pokreta studenata i formantne frekvencije (0.49); dok je mala korelacija bila slučaj kod formantne frekvencije i vremena govora (-0.21); pokreta studenata i pozicije najvećega autokorelacionog pika (0.39).

Generalno, korelacija između parametara pokazuje da je opadanje pažnje studenata praćeno povećanjem frekvencije pokreta u auditorijumu kao i povećanom manifestacijom energije, tona i vokalnog “pražnjenja” u zvučnom signalu koji predstavljaju glasan grupni govor, žamor; čime je potencijalno mnogo manje segmenata koji se mogu klasifikovati kao govor. Ovome svedoči i negativna korelacija između detektovanog govora i pokreta studenata u vrednosti od -0.57 kod neinteresantnih delova predavanja.

7. Evaluacija mašinskih metoda učenja

Pre implementacije sistema urađena je simulacija na setu prikupljenom za vreme predavanja koji je korišćen u prethodnom delu rada za analizu parametara i njihovu korelaciju. Cilj je da se odredi tačnost sistema pri prepoznavanju zainteresovanosti studenata upotrebom predloženih metoda, kao validaciona faza metode koja prethodi implementaciji.

Evaluacija različitih metoda učenja se uglavnom bazira na krosvalidaciji (Refaeilzadeh et al. 2008) , preciznosti i metrici tačnosti (Kotsiantis, 2007) (procentu tačnih predikcija podeljenih sa ukupnim brojem predikcija). U ovom istraživanju takođe je korišćena krosvalidacija na sledeći način: u svakoj iteraciji algoritam koristi k-1 deo dataseta da bi učio model koji se zatim koristi za predviđanje na delu dataseta koji nije korišćen za učenje.

7.1. Izbor klasifikatora

Na prethodno snimljenom i segmentiranom datasetu kako je objašnjeno u poglavlju 5 izvršena je evaluacija za veliki broj trenutno najkorišćenijih algoritama: Ada-boost M1, drva odlučivanja (eng. Random forest tree; Alternating decision tree), C.45, Bajesove mreže, metode potpornih vektora (eng. Support Vector Machines (SVM), Neuronske mreže (eng. Neural Networks (NN)), naivni bajes (eng. Naïve Bayes) i k-najbliži komšija (eng. K-nearest neighbors - KNN).

Evaluiran je set različitih često upotrebljavanih klasifikatora na datasetu prikupljenom na prethodno opisan način. Korišćeno je 14 parametara ekstraktovanih sa audio snimaka, jedan parametar sa slika i jedan sa akcelerometra. Ovi parametri su:

- prosečan broj glasovnih segmenata

- prosečan broj govornih segmenata
- delići govora
- glasovna stopa
- frakcije govora u vremenu
- prosečan broj krakih govornih segmenata
- formantna frekvencija
- pouzdanost formantne frekvencije
- spektralna entropija
- vrednost najvišeg vrha autokorelacije
- lokacija najvećeg vrha autokorelacije
- broj autokorelacionih pikova
- energija
- promena energije u vremenu
- deskriptori pokreta studenata
- deskriptori pokreta predavača

Dataset se sastoji od prosečnih vrednosti ovih parametara za deonice od 30 sekundi i korišćen je kao input za Weka toolkit koji u sebi ima implementaciju za sve prethodno opisane algoritme. Standardne postavke Weka alata su korišćene za klasifikaciju i rezultati su prikazani u tabeli ispod za 10 krosvalidacija.

Tabela 18 Karakteristike dataseta korišćenog u evaluaciji

	Grupa 1	Grupa 2	Grupa 3	Total
Broj deonica	46	40	35	121
Broj studenata	6	19	30	55
Prosečan broj ocena po deonici	5.4	17.6	27.1	16.7
Ukupan broj ocena po deonici	248	704	949	1901
Broj deonica označenih kao "interesantne"	32	28	23	83
Broj deonica označenih kao "neinteresantne"	14	12	12	38

Nedavna istraživanja su pokazala da dekompozicija greške klasifikatora omogućava značajan uvid u performance istog. Ova dekompozicija potiče od analize regresije (Sundmaecker et al. 2010), učenjem modela sa numeričkim izlaznim vrednostima.

Tabela 19 Rezultati krosvalidacije za različite klasifikatore

Algorithm	Group 1	Group 2	Group 3	Group 4	Total	Kappa
Adaboost M1(Freund & Schapire, 1996)	99.2	0.92	0.14	0.92	0.92	0.77
Random forest tree (Breiman, 2001)	88.9	0.89	0.3	0.88	0.89	0.64
Bayes Net	87.8	0.88	0.19	0.89	0.88	0.65
Alternating decision tree (Freund & Mason, 1999)	87.8	0.88	0.26	0.88	0.88	0.63
C4.5 decision tree (Quinlan, 1993)	87.8	0.88	0.34	0.87	0.88	0.59
SVM (Platt, 1998) (Keerthi et al. 2001)	83.3	0.83	0.55	0.82	0.82	0.37
Naïve Bayes (Aha & Kibler, 1991)	82.2	0.82	0.28	0.84	0.82	0.5
NN	82.2	0.82	0.36	0.82	0.82	0.47
KNN (John & Langley, 1995)	75.6	0.76	0.45	0.77	0.76	0.29

Generalno, algoritmi sa velikim varijansama: metode potpornih vektora, neuronske mreže, drva odlučivanja, uglavnom imaju mnogo bolje performanse pri klasifikaciji multidimenzionalnih konteksta.

Neuronske mreže i metode potpornih vektora sa značajnim uzorkom mogu postići maksimalnu tačnost pri klasifikaciji, dok je npr kod naivnog bajesa potreban relativno mali dataset. Jedan od nedostataka neuronskih mreža jeste što se pored odličnih performansi pri detekciji paterna ne može utvrditi šta je tačno uticalo na donošenje određenih rezultata.

U svakom slučaju, modeli sa veoma malim odstupanjem poput metode najbližeg komšije, bez obzira na kvalitet modela, nemogu imati visoke performanse ako im je uzorak ograničen (Manning, 2008).

Naivni bajes da ima visoka odstupanja zato što se pretpostavlja da se uzorak može sumirati pod jednom distribucijom verovatnoće i da je ovaj model dovoljan da uradi diskriminaciju klasa. U suprotnom, algoritmi sa visokim varijansama mogu generisati kompleksne modele koji bolje odgovaraju varijaciji podataka.

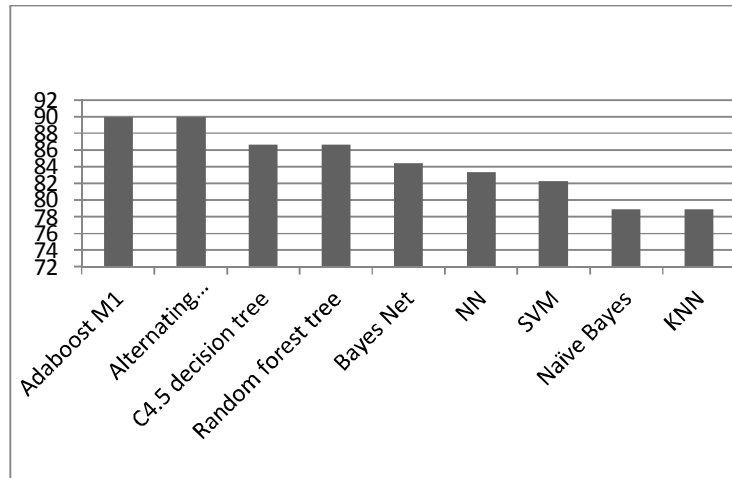
Primer algoritma sa visokim varijansama su drva odlučivanja, neuronske mreže i metode potpornih vektora. Očigledan nedostatak algoritama sa visokim varijansama je pre svega overfitting, odnosno pretreniranje.

Naivni bajes je robustan na vrednosti koje nedostaju pošto se one jednostavno ignorišu pri računanju verovatnoće i prema tome nemaju uticaj na proces klasifikacije.

S druge strane kNN i neuronske mreže zahtevaju kompletne podatke da bi performanse klasifikatora bile zadovoljavajuće.

U momentu kada je preliminarno testiranje rezultovalo zadovoljavajućim ishodom, klasifikator (mapiranjem neoznačenih klasa u klase) je spreman za upotrebu.

Klasifikatori koji imaju visoka odstupanja/male varijanse imaju prednost u odnosu na klasifikatore sa malim odstupanjima/velikim varijansama (Manning, 2008). Ovo se pokazalo kao tačno u slučaju Adaboost M1 i bajesovih mreža, koje su imale najbolje performanse u ovoj istraživanju: ova dva algoritma imala su najmanji broj netačnih pozitivnih klasifikacija (0.14 i 0.19). Međutim, naivni bajes je postigao bolje rezultate od drva odlučivanja i C4.5 algoritma. Najmanja tačnost zabeležena je kod KNN (75.6%), zatim neuronskih mreža (82.2%) i Naïve Bayes-a (82.2%). Metode potpornih vektora (83.3%) pokazale su se bolje od Naïve Bayes-a, neuronskih mreža i KNN, ali su imale veći broj netačnih pozitivnih klasifikacija. Konačno, najveću vrednost kapa koeficijenta imao je Adaboost M1 ($k = 0.77$) dok je najmanja zabeležena kod KNN ($kappa = 0.29$).



Slika 33 Tačnost klasifikatora pri krosvalidaciji nad dataset-om

Tačnost svakih evaluiranih algoritama prikazana je na slici iznad. Na osnovu dobijenih rezultata izabran je Adaboost M1 algoritam (Freund & Schapire, 1996) kao klasifikator za system pametne učionice.

7.2. Selekcija parametara

Da bi se evaluirao uticaj različitih parametara na tačnost klasifikatora sproveden je niz eksperimenata tako što je uklonjen jedan parametar za svaki proces klasifikacije na datom datasetu. Kao rezultat eksperimenta, dobijeni su rezultati prikazani u tabeli 18. Performanse klasifikatora su znatno opale uklanjanjem određenih parametara što je pokazatelj da su ovi parametri relevantni za istraživanje. Više od 14% opadanja performansi zabeleženo je kada su pokreti studenata i predavača uklonjeni iz seta podataka, kao i 6.7% kada spektralna entropija nije korišćena kao parametar. Kapa koeficijent je imao visoke vrednosti za spektralnu entropiju ($k = 0.54$), pouzdanost ($k = 0.69$), pokrete studenata ($k = 0.32$) i pokrete predavača ($k = 0.77$).

Mnogi parametri korišćeni u inicijalnoj klasifikaciji nisu bili relevantne za proces klasifikacije. Ovi parametri su uklonjeni iz seta podataka i krosvalidacija je ponovljena sa datasetom u kome su bili samo sledeći parametri: spektralna entropija, formantna frekvencija, pouzdanost u formantnu frekvenciju, predavačevi pokreti i pokreti studenata. Rezultati su bili nepromenjeni u odnosu na rezultate iz prethodnog poglavlja, prema čemu se može zaključiti da su dati parametri relevantni za istraživanje i da se upotrebom samo njihovog podskupa može doneti zaključak o zainteresovanosti studeata za vreme predavanja.

Tabela 20 Tačnost klasifikatora bez datih parametara

Parametar	Tačnost [%]	Kapa
Originalan set parametara	92.22	0.77
Bez pokreta studenata	80.00	0.38
Bez pokreta predavača	90.00	0.72
Bez promene energije u vremenu	92.22	0.77
Bez energije	92.22	0.77
Bez broja autokorelacionih vrhova	92.22	0.77
Bez lokacija autokorelacionih vrhova	92.22	0.77
Bez vrednosti najvišeg autokorelacionog vrha	92.22	0.77
Bez spektralne entropije	85.56	0.54
Bez pouzdanosti formantne frekvencije	90.00	0.69
Bez formantne frekvencije	92.22	0.77
Bez kratkih govornih segmenata	92.22	0.77
Bez delova govora u vremenu	92.22	0.77
Bez govorne stope	92.22	0.77
Bez prosečne dužine govornih segmenata	92.22	0.77
Bez prosečne dužine glasovnih segmenata	92.22	0.77

8. Dizajn i implementacija sistema za praćenje aktivnosti studenata

8.1. Modelovanje sistema za praćenja aktivnosti studenata

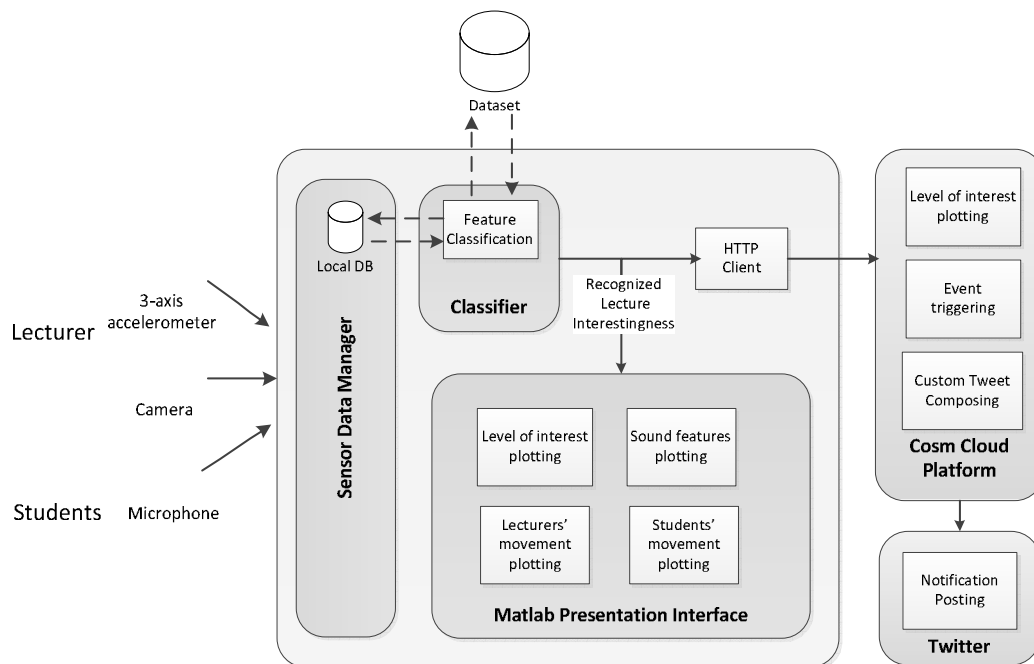
8.1.1. Opis sistema

Uloga sistema je prepoznavanje zainteresovanosti studenata za vreme predavanja kao relativne mere dobijene procesiranjem parametara iz zvuka i slike i njihovom klasifikacijom upotrebom nepovratnih neuralnih mreža. Sekvence slika zabeležene su kamerom, a zvuk mikrofonom u deonicama od 30 sekundi koje se direktno obrađuju upotrebom prethodno opisanih algoritama za obradu signala. Nakon obrade signala, pristupa se klasifikaciji korišćenjem neuralnih mreža, gde se za određene deonice iz kojih su prethodno izdvojeni glasovni segmenti, kao i segmenti bez govora, formantna frekvencija, frekvencija pokreta sa slike, itd, vrši prepoznavanje konteksta kao interesantnog ili neinteresantnog. Neuronske mreže svoj proces klasifikacije baziraju na klasifikatoru koji je prošao sve procesne faze, treniranja, validacije i testiranja.

8.1.2. Arhitektura sistema za praćenje aktivnosti studenata

Sistem je kompletno razvijen u softverskom paketu Matlab i sastoji se od sloja za prikupljanje podataka (prikupljanje podataka od kamere, mikrofona, acceleratora), sloja za analizu i obradu podataka i sloja za prezentaciju i skladištenje (slika 35). Struktura sistema je tako definisana da sloj za prikupljanje podataka koristi dve instance Matlab-a: jednu za snimanje zvuka i jednu za snimanje slike; dok sloj analize i obrade signala i sloj prezentacije i skladištenja dele jednu instancu. Sam proces prezentacije se takođe odvija u Matlab-u, ali se rezultati šalju do *Cosm* (Pachube) *cloud* platforme gde se omogućava skladištenje rezultata kao i prezentacija dostupna trećim licima. Notifikacije korisnicima omogućene su preko Tvitera.

Budući sistemi pametnih učionica u mnogome zavise od mogućnosti integracije različitih implementacija pametne učionice, tako da je neophodno razmotriti virtualne resurse i upotrebu cloud platformi koje bi obezbedile čuvanje podataka i interfejs za pristup rezultatima putem uvek dostupnih eksternih interfejsa. Prema tome relacioni model objekti-veze nije prikazan u ovom radu jel se kao krajnji rezultat klasifikacije dobije vrednost u vremenu koja se skladišti i grafički prikazuje preko *cloud* platforme.



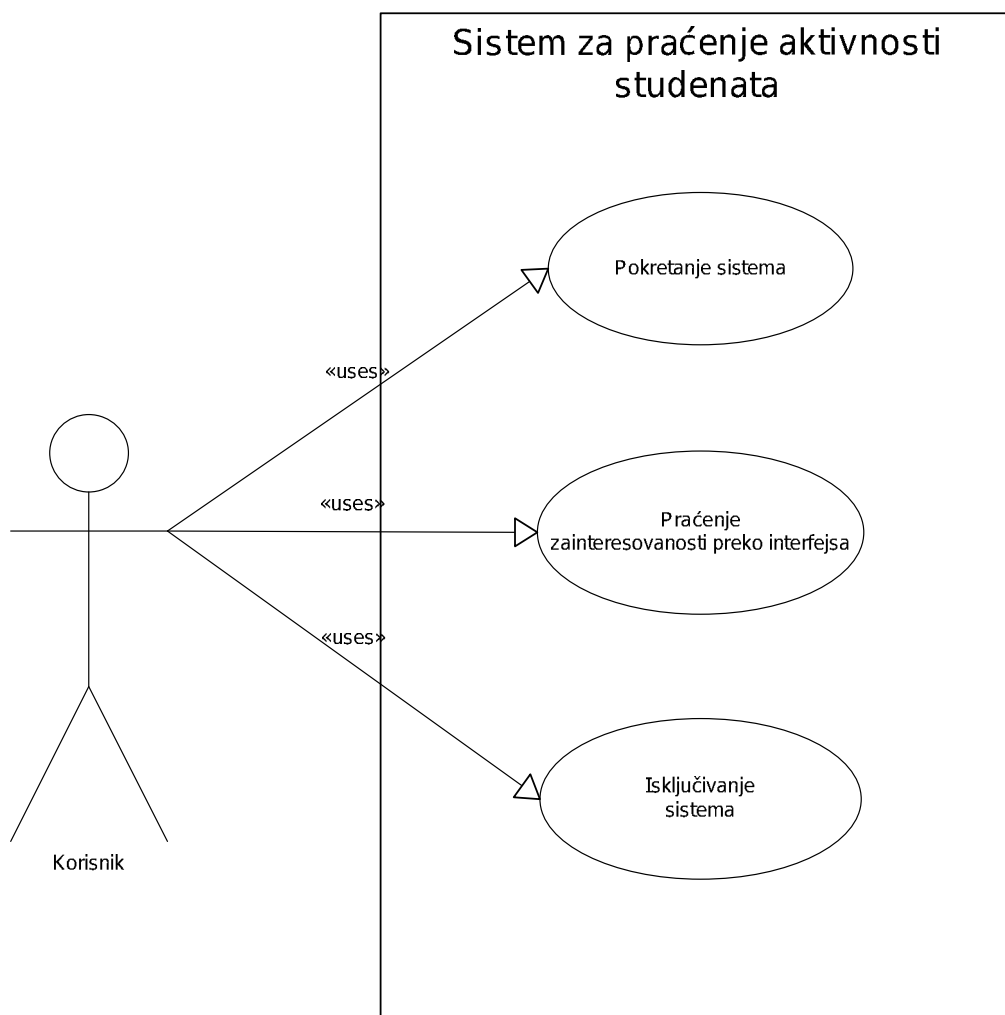
Slika 34 Arhitektura sistema

8.1.3. Dijagram slučajeva korišćenja

Svaki sistem ima predefinisane interakcije sa korisnikom ili automatizovanim akterom koji koristi sistem za određene procese za koje se očekuje da se izvršavaju predviđenim tokom. UML dijagrama slučajeva korišćenja se koristi za modelovanje ponašanja sistema ili delova sistema i predstavlja opis skupova akcija i varijanata koje sistem može da izvrši kako bi se dobio rezultat uočljiv za korisnika (Booch et al. 1999) .

Dijagram slučajeva korišćenja pokazuje skup slučajeva korišćenja, korisnika i njihovih korelacija.

Prema zahtevima sistema, dijagram slučajeva korišćenja sistema za praćenje aktivnosti studenata prikazan je na Slici 36:



Slika 35 Dijagram slučajeva korišćenja sistema

SK1: Slučaj korišćenja – Pokretanje sistema

Naziv: Pokretanje sistema

Svrha: Služi za pokretanje sistema za praćenje aktivnosti studenata

Akteri: Korisnik

Učesnici: Korisnik i sistem

Osnovni scenario:

1. Korisnik iz Matlab terminala pokreće sistem
2. Sistem prikazuje rezultate ekstrakcije

SK2: Slučaj korišćenja – Praćenje zainteresovanosti preko interfejsa

Naziv: Praćenje zainteresovanosti preko interfejsa

Svrha: Služi za praćenje zainteresovanosti studenata u realnom vremenu

Akteri: Korisnik

Osnovni scenario:

1. Korisnik iz Matlab terminala pokreće sistem
2. Sistem prikazuje rezultate ekstrakcije

SK2: Slučaj korišćenja – Praćenje zainteresovanosti preko interfejsa

Naziv: Praćenje zainteresovanosti preko interfejsa

Svrha: Služi za praćenje zainteresovanosti studenata u realnom vremenu

Akteri: Korisnik

Učesnici: Korisnik i sistem

Osnovni scenario:

1. Korisnik prati promenu zainteresovanosti u vremenu kao i promenu vrednosti parametara koji se ekstraktuju

SK2: Slučaj korišćenja – Isključivanje sistema

Naziv: Isključivanje sistema

Svrha: Služi za isključivanje sistema za praćenje zainteresovanosti studenata

Akteri: Korisnik

Učesnici: Korisnik i sistem

Osnovni scenario:

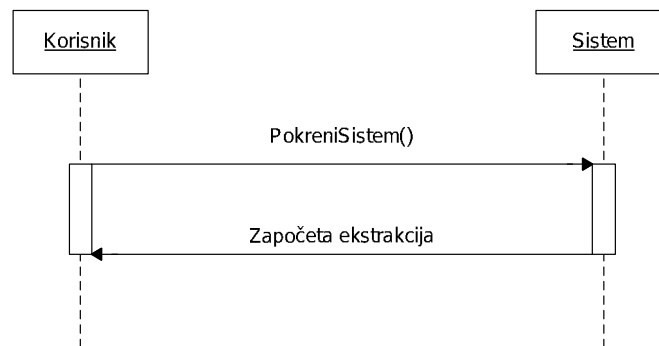
1. Korisnik zaustavlja rad sistema

8.1.4. Sistemski dijagram sekvenci

DS1: Dijagrami sekvenci slučaja korišćenja – Pokretanje sistema

Osnovni scenario:

1. Korisnik iz Matlab terminala pokreće sistem unošenjem komandi: *save_sound*, *save_screenshot*, *run_classroom*
2. Sistem vraća poruku da je ekstrakcija parametara započeta

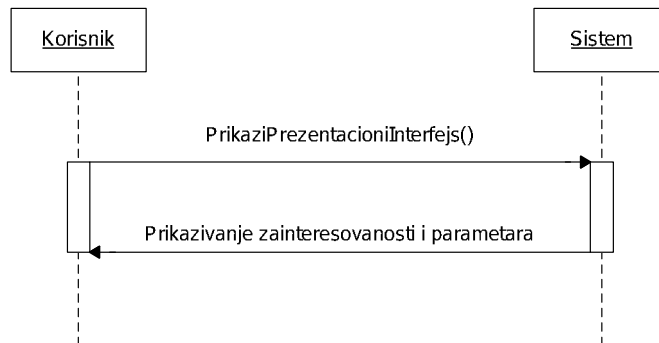


Slika 36 Dijagram sekvenci slučaja korišćenja – pokretanje sistema

DS2: Dijagrami sekvenci slučaja korišćenja – Praćenje zainteresovanosti preko interfejsa

Osnovni scenario:

1. Korisnik otvara prozor sa prezentacionim interfejsom
2. Sistem preko grafičkih dijagrama prikazuje trenutnu vrednost zainteresovanosti studenata, ukupnu vrednost zainteresovanosti studenata i vrednost ekstraktovanih parametara u deonicama.



Slika 37 Dijagram sekvenci slučaja korišćenja – Praćenje zainteresovanosti preko interfejsa

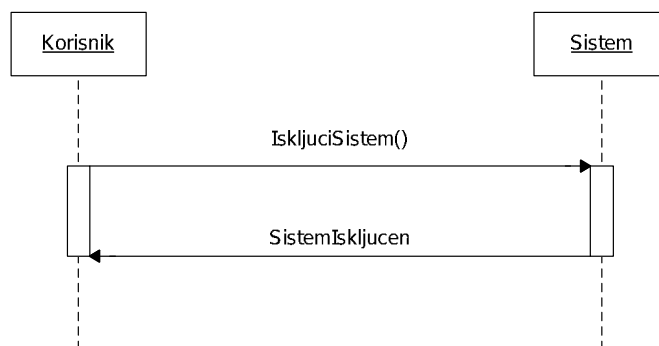
DS3: Dijagrami sekvenci slučaja korišćenja – Isključivanje sistema

Osnovni scenario:

1. Korisnik isključuje sistem za praćenje aktivnosti iz Matlab terminala unošenjem komande stop(t)

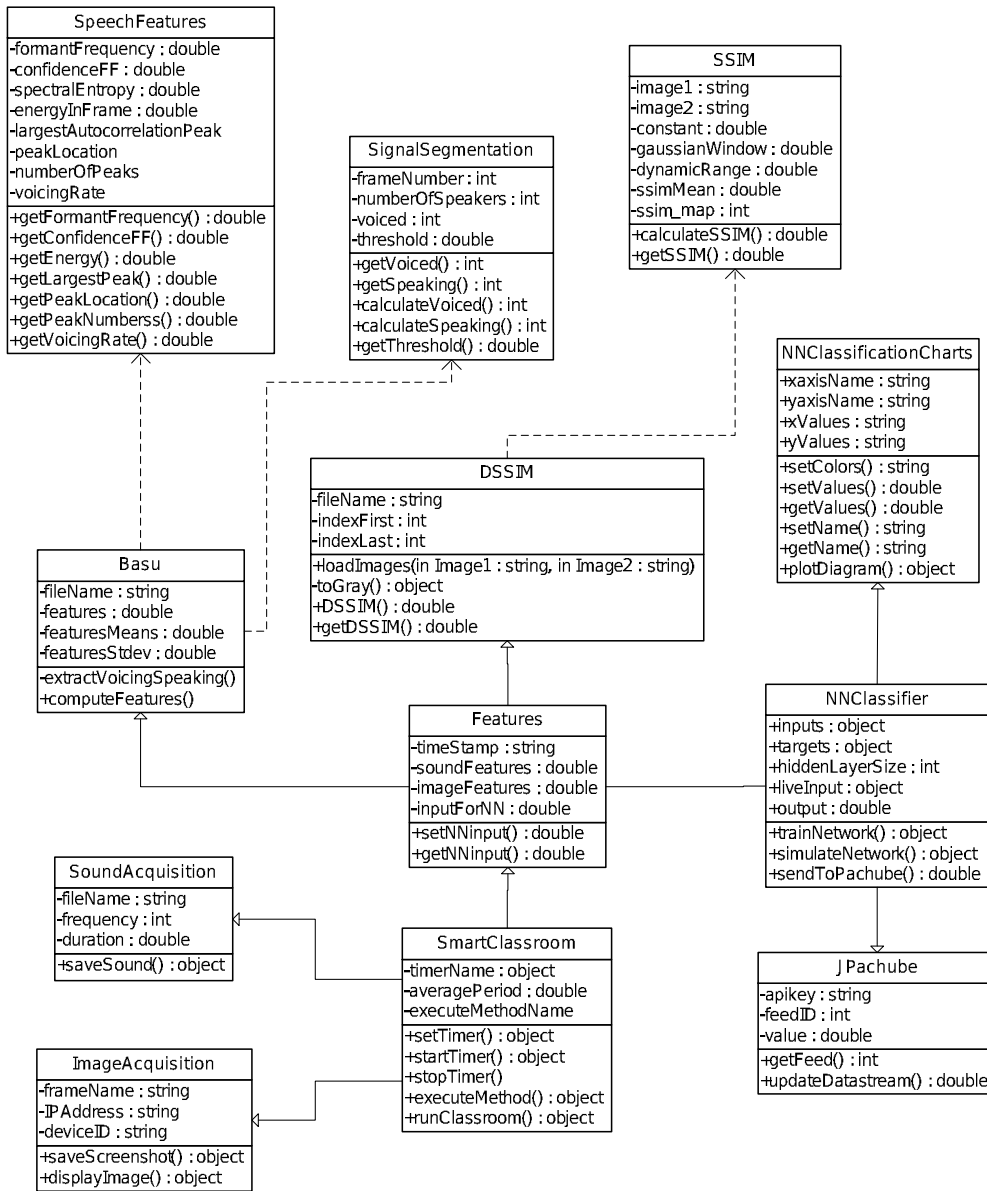
Alternativni scenario:

- 1.1 Ukoliko je sistem usled ekstrakcije parametara prvo će završiti ekstrakciju a nakon toga će prestati sa radom



Slika 38 Dijagram sekvenci slučaja korišćenja – Isključivanje sistema

8.1.5. UML diagram klasa



Slika 39 Dijagram klasa sistema

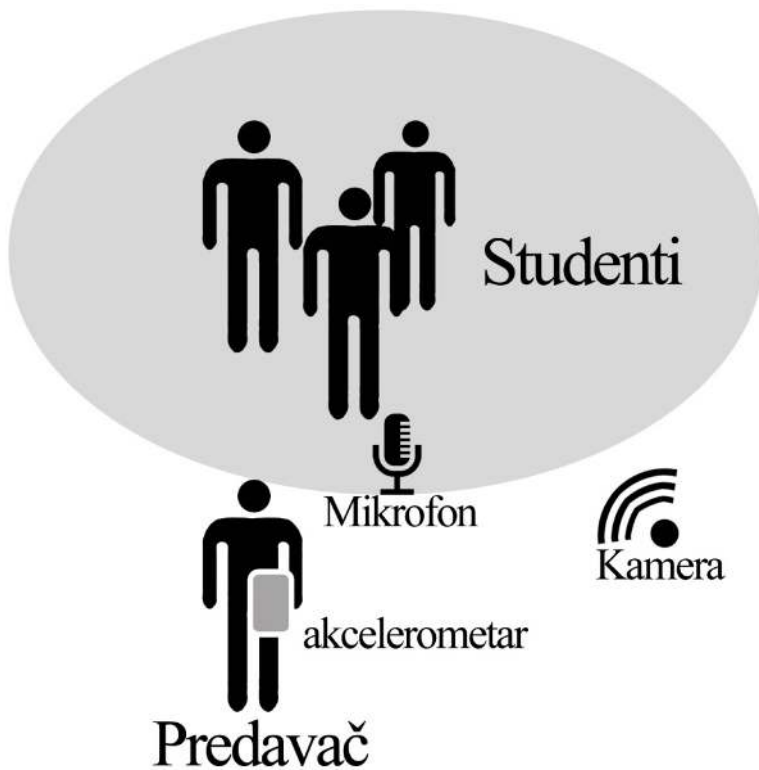
Dijagram predstavlja globalni opis systemske arhitekture, atributa i operacija unutar klasa koje prikazuju prethodno objasnjene funkcije unutar sistema.

8.2. Implementacija sistema za praćenje aktivnosti studenata

8.2.1. Sloj prikupljanja podataka

Postavka opreme u učionici je takva da se kamerom snimaju samo studenti kako bi se izbeglo uvrštavanje predavačevih podataka u set koji će biti korišćen kao input za klasifikaciju neuralnim mrežama. U ovoj postavci sistema mikrofon je direktno priključen na računar dok je IP kamera povezana preko WiFi mreže sa računarom.

Digitalni zapis zvuka čine prirodna mešavina ambijentalnog eha, kao i ostali uobičajenih zvukovi u učionici: žamora studenata, govora, ali i govora predavača. Ključno je da se na minimum svedu zvukovi koji nisu uobičajeni za predavanje, koji mogu da se pojave, npr. buka van učionice. Potrebno je isključiti mogućnost snimanja ovih zvukova jer značajno mogu da utiču na tačnost sistema.



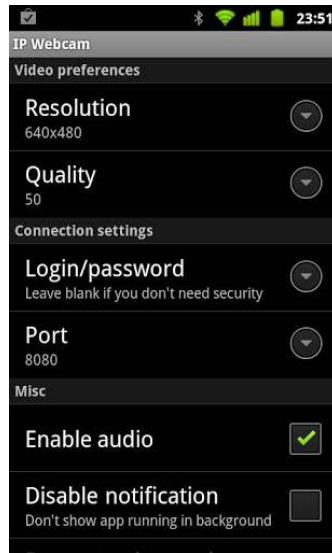
Slika 40 Prikupljanje podataka u učionici

Zvuk se snima na frekvenciji od 8KHz u deonicama od 30 sekundi i sačuvava u radnom prostoru Matlab-a. Neophodno je da frekvencija bude 8KHz jer algoritam za ekstrakciju osobina iz zvuka zahteva digitalni zapis u ovoj frekvenciji. Varijabla index se inkrementira kako bi se zvučni zapis u radnom prostoru mogao imenovati tako da se jedinstveno može odrediti deonica pri kasnijoj obradi. Klasifikator je treniran na stereo audio signalu tako da je potrebno da ulazni signal takođe bude stereo zbog čega se prvi kanal kopira tako da se dobiju dva kanala, odnosno stereo zvuk. Sličan pristup je upotrebljen pri treniranju klasifikatora (iako je zvuk stereo oba kanala sadrže isti zapis).

```
1  '----snimanje zvuka-----'
2  fs=8000;          % Sampling rate
3  duration=31;     % Recording duration
4  %index za ime fajla (zvuk)
5  index=1;
6
7  while(fs==8000)
8      y=wavrecord(duration*fs, fs);
9      fprintf('Deonica snimljena.\n');
10     %wavplay(y, fs);
11     fname = sprintf('sound%d.wav', index);
12     %snimanjem se dobija mono fajl. prvi kanal
13     %se kopira tako da se dobiju dva kanala
14     y = (y(:, [1 1]));
15     %snimanje u fajl
16     wavwrite(y, fs, fname);
17     index=index+1;
18 end
19
20 j=index;
```

Slika 41 Snimanje zvuka

Slika za potrebe ekstrakcije se ne snima kao video zapis, već se frejmovi direktno sačuvavaju u JPEG formatu u sekvencama od 1 sekunde. Kamera korišćena u ovom istraživanju je kamera Android telefona koja koristi aplikaciju IP Webcam (IP Webcam, 2013) sa interfejsom za snimanje screenshot-ova pozivanjem URL-a u lokalnoj mreži.



Slika 42 Aplikacija IP Webcam

Kada se kamera poveže na lokalnu mrežu moguće je definisati port za pristup resursu ukoliko neka druga aplikacija koristi port 8080 na lokalnoj adresi. Pristupom URL-u `http://dodeljenaadresa:8080/shot.jpg` aktivira se kamera koja snima frejm i direktno se omogućava pristupa snimljenom frejmu.

```

1  '-----snimanje frejmova-----'
2  %android kamera - snima frejmove u razmaku od 1 s
3  url = 'http://192.168.0.105:8080/shot.jpg';
4  ss = imread(url);
5  FileNum = 1;
6  fh = image(ss);
7  while(1)
8      pause(1)
9      ss = imread(url);
10     set(fh,'CData',ss);
11     drawnow;
12     fname = ['Image' num2str(FileNum) '.jpg'];
13     imwrite(ss, fname, 'JPG');
14     disp(['Sacuvana slika broj ' fname]);
15     FileNum = FileNum + 1;
16 end
17
18

```

Slika 43 Snimanje slike sa IP kamere

Sistem je moguće koristiti i sa kamerom povezanom direktno na računar ali je neophodno u tom slučaju koristiti drugi kôd (slika 45) pisan samo za proces ekstrakcije preko web/usb kamere. Običan aparat ili kamera nemaju fiksni režim snimanja, što može pri pravljenju frejmova direktno sa kamera da izazove kašnjenja koja nisu fiksna zbog čega se ne može raditi snimanje u zadatom intervalu. Ukoliko je kašnjenje veće - interval (linija 8 na slici slici 45) je potrebno izmeniti tako da snimanje slike bude tačno 1 sekund. Interval kašnjenja pre svega zavisi od hardverskih/softverskih karakteristika samog uređaja.

Kvalitet snimljene slike i zvuka neophodno je da bude što bolji ali i da postoji što manje razlika u kvalitetu i predispozicijama koje su korišćene u setu za klasifikaciju neuralnih mreža i setu koji se koristi za prepoznavanje parametara, odnosno u izvršnom okruženju sistema.

```
1 %snima frejmove sa web kamere
2 obj = videoinput('winvideo',1,'YUY2_640x480');
3 preview(obj);
4 start(obj);
5 FileNum = 0;
6
7 while(1)
8     pause(1)
9     data = getsnapshot(obj);
10    fname = ['Image' num2str(FileNum) '.jpg'];
11    imwrite(data, fname, 'JPG');
12    disp(['Sacuvana slika broj ' fname]);
13    FileNum = FileNum + 1;
14 end
```

Slika 44 Snimanje slike sa veb kamere

8.2.2. Sloj za obradu i analizu signala

Analiza signala obavlja se u Matlabu za deonice od 30 sekundi zbog čega postoji realno kašnjenje rezultata u istom intervalu ali nema dodatnih kašnjenja zbog procesiranja

signala iz prostog razloga što se procesiranje obavlja dok se naredna deonica snima. Svoj za obradu i analizu signala je implementiran u Matlab-u i koristi treću instancu Matlab-a koja u radnom prostoru ima sačuvane uzorke slike i zvuka u deonicama od 30 sekundi spremne za obradu. Po pokretanju sistema (komanda run_classroom) iz komadne linije Matlab-a, sistem prvi put ima odložen proces ekstrakcije za period od 30 sekundi kako bi se snimila deonica za isti period. Kao što je već naglašeno, nezavisno od procesa ekstrakcije, tajmer pokreće skriptu za ekstrakciju parametara na svakih 30 sekundi.

```
1 %ucitaj NN klasifikator
2 load Klasifikator;
3
4 %ucitaj java HTTP biblioteke
5 javaaddpath('C:\Program Files\MATLAB\R2011a\smartclassroom\lib\httpcore-4.2.2.jar');
6 javaaddpath('C:\Program Files\MATLAB\R2011a\smartclassroom\lib\httpClient-4.2.3.jar');
7
8 %prvih 30 frejmova (samo za prvi put)
9 N = 30;
10
11 %index za ime fajla (zvuk)
12 index=1;
13
14 %index za matricu zainteresovanost
15 j=1;
16
17
18 %sacekaj 30 sekund da se snime slika i zvuk
19 '----Cekanje da se deonice snime----'
20 pause(30);
21
22 %definisaj tajmer koji ce se koristiti da pokrece ekstrakciju na svakih 30 sekundi
23 '----Definisanje tajmera----'
24 t = timer;
25 set(t,'executionMode','fixedRate');
26 set(t,'period',30);
27 set(t,'TimerFcn','extract_features');
28 start(t);
```

Slika 45 Pokretanje pametne učionice

```

1  '----Ekstrakcija DSSIM-a zapoceta-----'
2  dssim;
3
4  '----Ekstrakcija DSSIM-a završena-----'
5
6  '----Ekstrakcija osobina iz zvuka zapoceta-----'
7  sound_name = sprintf('sound%d.wav', index);
8  features = speech_features_stereo(sound_name);
9  [states_voiced, states_speaking] = voicing_speaking(features , 'threshold');
10 [means, stds, others] = chunk_features(features, states_voiced, states_speaking, 0.5);
11 '----Ekstrakcija osobina iz zvuka za period završena-----'
12
13 '----Spreman input za klasifikaciju-----'
14 liveinput(1:6) = others(1:6);
15 liveinput(7:14) = means(1:8);
16 liveinput(15:22) = stds(1:8);
17 liveinput(24) = DSSIMmean(1);
18
19 '----Klasifikacija zapoceta-----'
20 zainteresovanost(j)= sim(net, liveinput. ');
21 '----Klasifikacija završena-----'
22
23 %Prikazi dijagram sa karakteristikama
24 dijagram_karakteristike;
25
26 '----Update-ujem Cosm feed-----'
27 post_to_pachube;
28
29 %inkrementiraj index za matricu
30 j=j+1;
31
32 '----Cekam 30 sekundi-----'
33

```

Slika 46 Ekstrakciju parametara

```

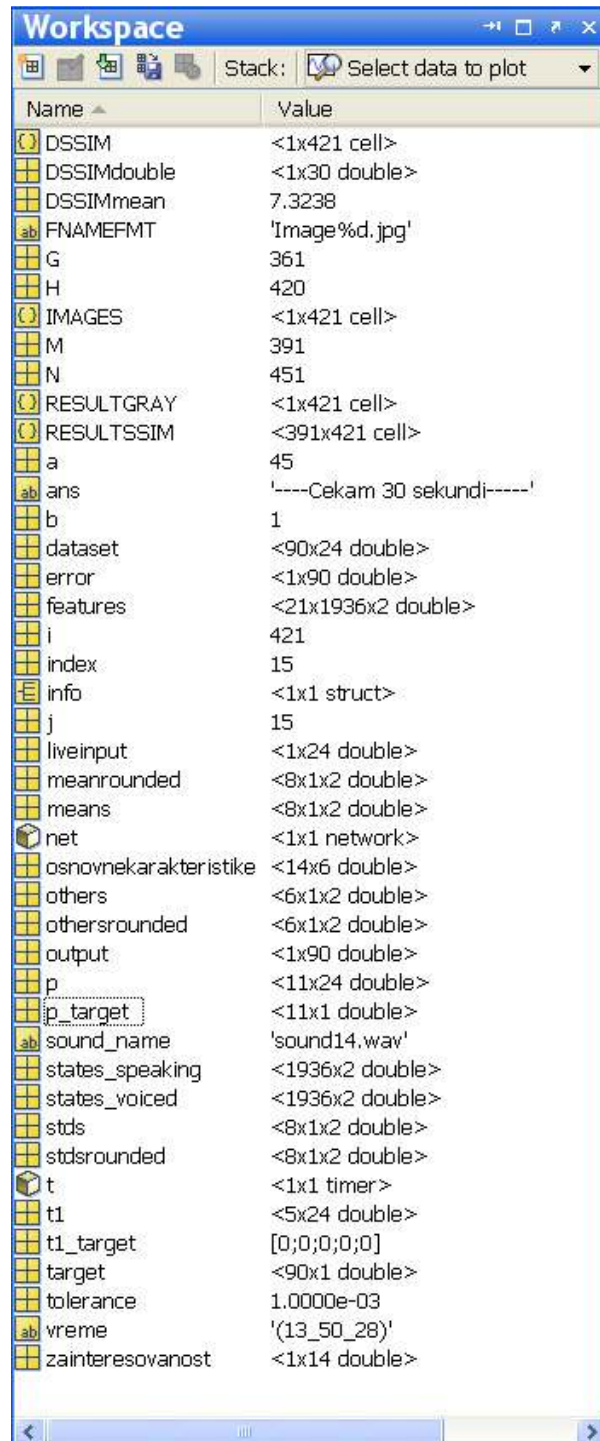
1  %*****
2  %DSSIM INDEX poredjenje za slike
3  %*****
4
5  %file = dir('*.jpg');
6  %N = length(file);
7
8  M=N-30;
9  IMAGES = cell(1,N);
10 FNAMEFMT = 'Image%d.jpg';
11
12 % Load images
13 for i=M:N
14     IMAGES{i} = imread(sprintf(FNAMEFMT, i));
15 end
16
17 % Prebaci u gray
18 RESULTGRAY = cell(1,N);
19 for i=M:N
20     RESULTGRAY{i} = rgb2gray(IMAGES{i});
21 end
22
23 %izracunaj SSIM
24 RESULTSSIM = cell(M,N);
25 for i=M:N
26     if i==N
27         break
28     else
29         RESULTSSIM{i} = ssim_index(RESULTGRAY{i},RESULTGRAY{i+1});
30     end
31 end
32
33 DSSIM = cell(1,N);
34 for i=M:N
35     if i==N
36         break
37     else
38         DSSIM{i} = (1/(1-ssim_index(RESULTGRAY{i},RESULTGRAY{i+1})));
39     end
40 end
41
42 %izracunati prosecan DSSIM za deonicu
43 DSSIMdouble = cell2mat(DSSIM(M:(N-1)));
44 DSSIMmean = mean(DSSIMdouble);
45
46
47 N=N+30;
48

```

Slika 47 Ekstrakcija DSSIM index-a između frejmova

Na izlazu, neuralne mreže dodeljuju vrednost između 0 i 1, što predstavlja relativnu meru zainteresovanosti: što je vrednost bliža 1, studenti su zainteresovaniji i obrnuto, što je bliža 0 veća je verovatnoća da je zainteresovanost u određenoj deonici manja. Rezultate analiziranih deonica moguće je pratiti putem prezentacionog interfejsa. Gotovi rezultati dobijeni nakon klasifikacije neuronskih mreža, prikazuju se na dijagramu i istovremeno šalju do Cosm cloud platforme koja dalje omogućava prezentaciju aktivnosti nezavisno od mesta na kome se korisnik koji prati rezultate nalazi.

Na slici 49 prikazani su svi objekti, varijable i matrice koje se koriste u sistemu.



Slika 48 Radni prostor Matlab-a

8.2.3. Sloj za prezentaciju i skladištenje

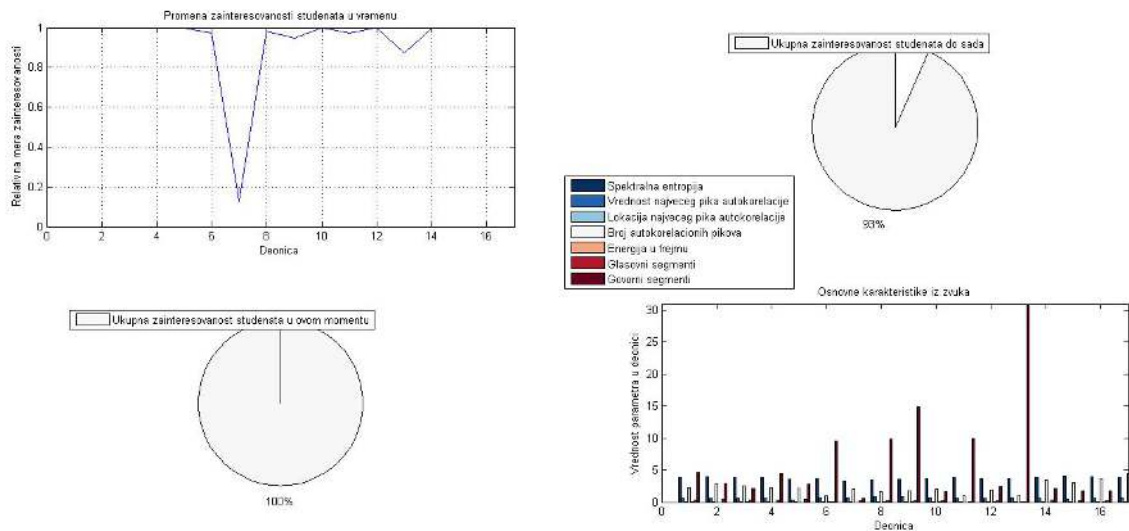
Prezentacija podataka se obavlja pomoću lokalnog interfejsa u Matlabu i preko eksterne Cosm platforme koja ujedno skladišti podatke u cloud-u. Preko lokalnog interfejsa predavač može da prati zainteresovanost studenata i vrednosti parametara, dok se eksterni može koristiti za pristup od strane drugih platformi, učionica, itd.

8.2.3.1. Lokalni prezentacioni interfejs

Ekstraktovani parametri kao i krajnji rezultat klasifikacije se prikazuju na četiri dijagrama:

1. Dijagram promene zainteresovanosti u vremenu
2. Dijagram ukupne zainteresovanosti studenata od početka predavanja
3. Dijagram ukupne zainteresovanosti studenata u datom vremenu
4. Dijagram sa osnovnim karakteristikama iz zvuka

Svaki od navedenih dijagrama se menja u realnom vremenu, prikazujući relevantne rezultate.



Slika 49 Lokalni prezentacioni interfejs

8.2.3.2. Eksterni interfejs za skladištenje i prezentaciju

Cloud platforma Cosm je korišćena za vizuelizaciju i omogućava pristup merenjima trećoj strani. Svaku učionicu moguće je registrovati kao jedan kanal na Cosm servisu, odakle se dalje podaci mogu preuzimati i koristiti za analizu i deljenje. Svi podaci sačuvavaju se u Cloud-u, tako da upotreba baze podataka nije neophodna.

Cosm je klad platforma za Internet inteligentnih uređaja koja u realnom vremenu obezbeđuje funkcionalnosti upotrebom aplikacionog programskog interfejsa (API). Posredstvom API-a IoT uređaji se jednostavno konektuju na patformu, gde se dalje merenja u realnom vremenu mogu analizirati, pregledati istorijat predhodnih merenja, kao i upotrebljavati alati za upravljanje senzorima, uređajima i okruženjima koji se mogu nalaziti bilo gde u svetu. Sastavni delovi platforme su okruženje koje definiše mereni entitet, strim podatka i vrednost u strimu.

Entitet

Entitet je određen skup podataka koji je kontekstualno povezan i najčešće se nalazi na određenoj geolokaciji. Pod entitetom se podrazumeva bilo fizički entitet poput jednog uređaja, prostorije, zgrade ili oblasti; ili virtuelni entitet koji može da podrazumeva monitoring protoka servera, ili bilo kakav virtuelan skup vrednosti koji je statičan ili se menja u vremenu. Cosm feed je reprezentacija samog entiteta pomoću podataka. Entitet pametne učionice se registruje kao jedan zaseban feed čime je obezbeđena dalja vizualizacija podataka.

Strim podataka

Strim podataka je reprezentacija individualnog merenja senzora u određenom okruženju. Svaki strim podataka mora da ima jedinstveni alfranumerički ID. Jedinica podataka se takođe može definisati kao i novi tagovi. Vrednost u strimu je vrednost u određenom vremenu i sastoji se od para: vrednost-vreme.

Podržani formati u okviru platforme su JSON, CSV, XML, itd, tako da sam izbor formata podataka zavisi od strane korisnika. Tip podataka JSON je efikasan za veb aplikacije, pregledan je i lako se može parsirati. Takođe ima mnogo manji overhead od XML-a, tako da troši znatno manje protoka.

Strimovi podataka se šalju u formatu „vrednost-vreme“, gde vrednost može biti između 0 i 1, dok je vreme običan timestamp u momentu slanja podataka do platforme za vizuelizaciju.

HTTP i REST (Representational State Transfer) interfejs

Interfejs platforme kompatibilan je sa HTTP zahtevima i podleže REST (Fielding, 2000) principima. Tri glavna tipa podataka (entitet, strim i vrednost), hijerarhijski su predstavljeni u URI-u na način koji odražava njihovu zavisnost: entitet – strim – vrednost. REST pristup ovim tipovima je moguć upotrebom različitim metoda koji određuje i karakteristiku akcije na određenom tipu:

GET : Vraća trenutno stanje objekta PUT : Setuje stanje objekta na određenu vrednost POST : Kreira novi objekat DELETE : Briše objekat

Za svaki od navedenih zahteva dobija se kao odgovor jedan od standardnih HTTP statusa:

200 OK: zahtev uspešno procesiran 401 Not Authorized: autentifikacije nije uspešno prošla 403 Forbidden: Platforma je uspešno dobila zahtev ali odbija da ga ispuni. Detalji su uključeni u nastavku svakog zahteva 404 Not Found: Zahtevan je nevalidan URI ili resurs ne postoji 422 Unprocessable Entity: Feed nije uspešno na pravljen, format podataka nije validan (JSON)

503 No server error: Previše zahteva poslatih od strane klijenta. Server ne može da ih obradi.

Nisu svi HTTP klijenti u mogućnosti da pošalju PUT ili DELETE zahtev (na primer veb brauzeri). Iz tog razloga da bi se izbegle bilo kakve nekompatibilnosti, PUT zahtev se može poslati kao POST zahtev sa "_method" parametrom, npr.:

POST na URI `http://api.cosm.com/v2/feeds/87441/datastreams/0.csv?_method=put` simuliraće PUT

POST na URI
`http://api.cosm.com/v2/feeds/87441/datastreams/0.csv?_method=delete` simuliraće
DELETE

Bezbednost - SSL/HTTPS

Kao autentifikaciona metoda mogu se koristiti HTTPS i SSL tako što se jednostavno zameni HTTP sa HTTPS u zahtevu.

Nalog i dostupne funkcionalnosti

Cosm platforma je besplatna sa ograničenjima po pitanju broja zahteva koje je moguće poslati u određenom vremenskom periodu. Broj zahteva koje je moguće poslati u jednoj minuti je 100, što je sasvim dovoljno za potrebe predloženog sistema.

Pregled interfejsa

Slanje podataka iz Matlaba obavlja se upotrebom Java metoda i Apache HTTP klijenta koji je importovan u Matlab okruženje. Dijagram prikazuje relativnu meru zainteresovanosti kao rezultat klasifikacije neuronskih mreža koje na izlazu daju vrednost između 0 i 1.



Slika 50 Pachube prezentacioni interfejs

Preko HTTP POST metode sa parametrom PUT u URL-u kao alternativa za PUT zahtev, merenja se šalju na Cosm servis kao CSV vrednost za svaku klasifikovanu deonicu od 30 sekundi. Autentifikacija se obavlja dodavanjem API ključa u heder paketa tako da je fid moguće ažurirati samo ako je ključ validan.

```

1  javaaddpath('C:\Program Files\MATLAB\R2011a\smartclassroom\lib\httpcore-4.2.2.jar');
2  javaaddpath('C:\Program Files\MATLAB\R2011a\smartclassroom\lib\httpclient-4.2.3.jar');
3
4
5  import org.apache.http.impl.client.DefaultHttpClient;
6  import org.apache.http.client.methods.HttpPost;
7  import org.apache.http.entity.StringEntity;
8
9
10 httpclient = DefaultHttpClient();
11
12 httpost = HttpPost('http://api.cosm.com/v2/feeds/87441/datastreams/0.csv?_method=put');
13 httpost.addHeader('Content-Type','text/plain');
14 httpost.addHeader('X-APIKey','S_HRKHsELey7N0gob4V1Oqqg5QvB0VoJwsGCD8ESYsw');
15
16 params = StringEntity(num2str(zainteresovanost(j-1)));
17 httpost.setEntity(params);
18
19 response = httpclient.execute(httpost);

```

Slika 51 Java metode koje se koristi za slanje podataka iz Matlaba do Cosm fida

8.2.3.3. Tviter notifikacije

Cosm platforma je povezana sa *Twitter* nalogom preko API-a koji omogućava slanje kastomiziranih obaveštenja koje studenti, korisnici i treća lica mogu da prate direktno preko socijalnih mreža. Notifikacije su postavljene tako da kada Cosm fid postane

aktivan posle dužeg vremena, odnosno kada je predavanje započeto, ili kada fid ne prima nova ažuriranja, odnosno kada je predavanje završeno. Takođe ukoliko je predavanje interesantno ili ne korisnici mogu takođe pratiti preko *twitter*-a.



Slika 52 Obaveštenja putem Tviter-a

Na ovaj način omogućava se ukoliko je u toku konferencija ili veći broj predavanja istovremeno, da studenti odluče koje predavanje žele da prate. Za svaku učionicu se može napraviti jedan Cosm fid koji će zatim na isti Tviter nalog da šalje obaveštenja tako da korisnici ostanu informisani za sve sale u kojima se održava predavanje.

9. Evaluacija performansi sistema u realnom okruženju

U ovom poglavlju prikazani su rezultati evaluacije sistema u realnom okruženju. Sistem je evaluiran u učionici snimanjem dodatnih serija predavanja. Izlazna vrednost klasifikacije sistema je upoređena sa ocenama studenata za vreme predavanja. Pošto je cilj da razvijen sistem bude robustan na broj studenata i na različite predavače, algoritam je testiran na tri grupe studenata:

- 1. grupa od 6 studenata
- 2. grupa od 19 studenata
- 3. grupa od 30 studenata

Predavanja su držali različiti predavači u različitim učionicama od predavača i učionica u datasetu. Nisu sve ocene studenata uzete u obzir: glasovi su segmentirani upotrebom metodologije prethodno upotrebljenje za anotaciju podataka za dataset: glasovi (ocene) su segmentirane tako što su uzimane u obzir samo deonice gde su najmanje 90% studenata prisutnih predavanju glasali. Ova metoda je korišćenja zbog mogućnosti da broj glasova u nekim deonicama bude vrlo mali.

Ukupna dužina od 14.2 sati predavanja nakon segmentacije svedena je na 60.5 minuta (121 deonicu od 30 sekundi) dataseta. Grupe 1, 2 i 3 imale su prosečno po 5.4, 17.6 i 27.1 glasovi po deonici, ukupno 1078 glasova za sekvence ocenjene kao interesantne i 823 ocenjene kao ne interesantne od strane studenata. Više detalja o datasetu prikazano je u tabeli 19.

Tabela 21 Karakteristike dataseta korišćenog za evaluaciju

	Grupa 1	Grupa 2	Grupa 3	Ukupno
Broj deonica	46	40	35	121
Broj studenata	6	19	30	55
Prosečan broj ocena po deonici	5.4	17.6	27.1	16.7
Ukupan broj ocean po deonici	248	704	949	1901
Broj deonica ocenjenih kao interesantni	32	28	23	83

Broj deonica ocenjenih kao neinteresantni	14	12	12	38
-------------------------------------------	----	----	----	----

Da bi se dobili statistički značajni rezultati, korišćeno je 10 iteracija – svaka deonica je 10 puta evaluirana sa klasifikatorom i uzete su samo srednje vrednosti evaluacije za svaku deonicu. U tabeli su prikazani srednje vrednosti rezultata evaluacije sistema.

Kao što se može videti iz tabele, predložen algoritam je postigao dobre rezultate za sve tri grupe studenata: 83.44%, 81.2%, 81.1%. Kapa koeficijent je bio najviši za grupu od 6 studenata i opadao je sa povećanjem broja studenata. Prosečna vrednost tačnih pozitivnih klasifikacija bila je 0.81. Preciznost klasifikatora (0.84, 0.81, 0.78) je zadovoljavajuća kada se uzmu u obzir značajne razlike između okruženja u kom je evaluacija urađena i trenažnog okruženja. Broj lažnih pozitivnih klasifikacija nije zanemarljiv i vrednost od 0.42 za sve grupe pokazuje da sistem može u određenim situacijama pogrešno da detektuje zainteresovanost studenata. Povećavanje broja sekvenci za koje se vrši klasifikacija i uzimanjem srednje vrednosti u obzir (npr. za period od 5 minuta) povećaće tačnost sistema za posmatrani period.

Tabela 22 Performanse klasifikatora u učionici

Grupa	Tačnost (%)	TP	FP	Preciznost	Kapa
Grupa 3	81.01	0.80	0.47	0.78	0.39
Grupa 2	81.19	0.81	0.37	0.81	0.41
Grupa 1	83.44	0.83	0.42	0.84	0.47

Rezultati evaluacije pokazuju da se može postići prosečna tačnost od 81.88% identifikacije zainteresovanosti studenata za vreme predavanja za dataset nekorišćen u procesu treniranja. Ovo su prvi rezultati koji pokazuju kako da se oceni zainteresovanost studenata za vreme predavanja u učionici upotrebom metoda mašinskog učenja. Pervazivan pristup, metode korišćene i postignuti rezultati u ovom istraživanju otvaraju nove mogućnosti za nova istraživanja i aplikacije.

9.1. Analiza performansi klasifikatora i rezultata istraživanja

Praćenje sociološke aktivnosti i detekcija ljudskog ponašanja upotrebom tehnologije je jedan od najzahtevnijih procesa zbog velikih broja faktora koji utiču i na najjednostavnije slučajeve prepoznavanja. U ovom istraživanju, dizajniran je model baziran na algoritmu za mašinsko učenje koji može da razlikuje interesantno od neinteresantnog predavanja. Performanse sistema su evaluirane na određenom uzorku predavanja da bi se utvrdila tačnost sistema.

Način na koji je sprovedeno ovo istraživanje i poznati konteksti određeni na osnovu ocena predavanja koje su studenti izrazili u realnom vremenu, omogućavaju detaljniju analizu nelinearnog sistema.

Kao krajni cilj postavljeno je razvijanje sistema koje je u mogućnosti da razlikuje između interesantnih i neinteresantnih predavanja, robustan na broj studenata koji prati predavanje i na razlike koje se mogu javiti kod zbog diverzije tona između glasova različitih predavača, evaluirane su performanse na tri nove grupe studenata kojima je drugi predavač držao predavanje. Sistem je u ovom slučaju bio u mogućnosti da prepozna pažnju studenata sa tačnošću od 83.44%, 81.2% i 81.1% za svaku od grupa.

U ovom istraživanju, pri analizi socioloških signala iz zvuka i pokreta, pokazano je da spektralna entropija, formantna frekvencija i pokreti studenata igraju ključnu ulogu pri određivanju pažnje studenata.

Rezultati ekstrakcije parametara i njihove korelacije, pokazuju da je opadanje pažnje studenata praćeno povećanjem frekvencije pokreta (75% u ovom istraživanju) u auditorijumu kao i povećanom manifestacijom energije, tona i vokalnog "pražnjenja" u zvučnom signalu koji predstavljaju glasan grupni govor, žamor; čime je potencijalno mnogo manje segmenata koji se mogu klasifikovati kao glasovi i govor. Ovome svedoči i negativna korelacija između detektovanog govora i pokreta studenata u vrednosti od -0.57 kod neinteresantnih delova predavanja.

Osobine autokorelacije iz signala su ekstraktovane pre svega zbog važnosti ove karakteristike pri određivanju glasovnih segmenata i segmenata bez glasova. Glasovni segmenti koji imaju veći broj autokorelacionih pikova, prema čemu je najveći autokorelacioni pik kao i broj pikova jasan indikator o vrsti segmenta. U psihologiji stres govornika se manifestuje kao dominantno ponašanje koje reflektuje koliko dugo kao i koliko glasno osoba govori (Dunbar & Burgoon, 2005). Pošto studenti uglavnom pasivni receptori informacija, diskusija za vreme predavanja je više efektna i može da ima pozitivan uticaj na pažnju studenata (Steinert & Snell, 1999). Individualno glasovni segmenti i aktivnost iz grupnog signala je teško modelovati kada je mnogo osoba prisutno na snimku. U okruženju kao što je učionica, pretpostavka je da predavač priča većinu vremena i da se povremeno studenti uključuju u interakciju. Za vreme interakcije ostatak auditorijuma ako sluša praviće manje buke, tako da je moguće detektovati više govornih segmenata i prepoznata govorna aktivnost i glasovni segmenti bi trebalo da su visoki. U suprotnom, kada studenti ne prate predavanje, manje je segmenata koji mogu biti detektovani kao govor (manja je autokorelacija u signalu) i više je buke i šuma u signalu (žamora). Ovo je podržano činjenicom da je za interesantne delove predavanja, detektovanih govornih segmenata bilo 68% više.

Korelacija između parametara pokazuje da je opadanje zainteresovanosti studenata u slučaju interaktivnih predavanja praćeno sa povećanom frekvencijom pokreta, pričanjem sa kolegama, gledanjem okolo, igranjem sa predmetima, itd.

10. Zaključak

Zbog multidisciplinarnosti procesa praćenja aktivnosti studenata, do sada nije bilo radova koji su se sa metodološkog aspekta, kao i sa aspekta implementacije bavili ovom temom. Sistem za praćenje socioloških aktivnosti predstavlja jedinstven skup kompleksnih metoda za obradu digitalnog signala koje zajedno rezultuju relevantnim informacijama. Glavna hipoteza razvijena i dokazana u okviru ove teze je da se primenom Interneta inteligentnih uređaja u nastavi može poboljšati efikasnost nastavnog procesa kroz realizaciju sistema za praćenje aktivnosti studenata, koji omogućava predavaču da za vreme predavanja ima uvid u kvalitet predavanja kako bi mogao da utiče na isti.

U disertaciji predstavljena je detaljna analiza problema i identifikovan je skup parametara koji je u skladu sa postavljenim zahtevima sistema. Dati parametri su validirani eksperimentalnim merenjima i kroz anketu sprovedenu na uzorku od 300 studenata, a kasnije predloženi parametri su ekstraktovani iz signala prikupljenog za vreme predavanja, napravljena je njihova korelacija i na osnovu rezultata urađena je simulacija sistema, koja je pokazala 92.2% tačnosti pri određivanju zainteresovanosti studenata za vreme predavanja. Pokazano je da su najznačajniji parametri za detekciju pažnje spektralna entropija i formantna frekvencija iz zvuka kao i intezitet pokreta predavača.

Nakon procesa simulacije, sistem je implementiran i dodatna evaluacija pokazala je zadovoljavajuću tačnost (81.9%) nastalu kao rezultat upoređivanja ocena studenata koji su u realnom vremenu ocenjivali predavanje pomoću aplikacije i merenja koje je pokazao sam sistem.

Rezultati istraživanja saopšteni su na više skupova u zemlji i inostranstvu i objavljeni su u naučnom časopisu od međunarodnog značaja.

10.1. Kritički osvrt na sprovedeno istraživanje

Na kvalitet predavanja, koji predstavlja centralni fokus ovog istraživanja, utiče mnogo različitih faktora od različitih metoda podučavanja (Morgan & Kingston) do percepcije procesa učenja od strane studenata (Sharma, 1997). Cilj ovog istraživanja je da se odredi kvalitet predavanja kao mera zadovoljstva studenata upotrebom sistema koji prikuplja i klasifikuje parametre upotrebom tehnologije Interneta inteligentnih uređaja i metoda mašinskog učenja.

Razlike u rezultatima za drugi uzorak koji je bio potpuno različit od onog korišćenog za vreme procesa treniranja i validacije ukazuje da su tačnost, senzibilnost i specifičnost svakog porepoznavanja u strogoj povezanosti sa distribucijom zvuka u prostoriji, brojem ljudi prisutnih u učionici; kao i da je sistem podložan kulturološkim uticajima gde biheviroške razlike utiču na to da se drugačije kanališu određena stanja svesti kroz verbalne i gestikulacione kanale. Međutim generalno upotreba algoritma za mašinsko učenje je pristup koji omogućava da se ovakvi konteksti identifikuju, ali je bitan način prikupljanja, obrade, kao i sama konstrukcija parametara koji su korišćeni u istraživanju. Rezultati iniciraju da predložen sistem može da omogući prevazivan pristup za detekciju pažnje studenata, a kao najveća prepreka koja može uticati na tačnost sistema, izdvajaju se potencijalne razlike između trenažnog i izvršnog okruženja (drugačiji broj studenata, tip predavanja, razlika u akustici učionice, itd). Međutim, pored navedenih problema, korelacija identifikovana u ovom istraživanju, pruža univerzalnu metodu za identifikaciju datih konteksta, ne samo u ovom konkretnom primeru interaktivnog predavanja, već i u drugim primenama uopšte.

Tačnost sistema naknadno je moguće povećati dopunjavanjem parametara koji se mogu meriti u učionici. Kao krajnji rezultat, predavaču se obezbeđuje u gotovo realnom vremenu da stekne uvid u stepen zainteresovanosti studenata, što mu omogućava da zna u kom momentu je potrebno da određenu temu obradi u više detalja, promeni način

izlaganja ili da primeni neke druge metode podučavanja koje će da mu pomognu da se pažnja auditorijuma održi na visokom nivou.

10.2. Budući pravci istraživanja

Pored parametara evaluiranih u ovom istraživanju, pretpostavka je da u okruženju učionice ima još relevantnih parametara koji se mogu analizirati. Prema tome, dalje istraživanje biće usmereno u ovom pravcu kako bi se dodatni IoT uređaji uključili u sistem radi analize i poboljšanja performansi sistema.

Algoritmi za mašinsko učenje su evaluirani upotrebom standardnih postavki unutar Weka alata. Pošto sama evaluacija svih metoda prevazilazi obim ove studije, kao deo budućeg istraživanja, biće urađena evaluacija svih algoritama za sve dostupne varijacije postavki za odgovarajući algoritam kako bi se utvrdilo da li parametarizacija može da utiče i da da različite rezultate i kako bi se pokazalo koji je algoritam efikasniji od ostalih.

Dalji rad u ovoj oblasti je moguće usmeriti ka integraciji predloženog rešenja sa sistemom pametne učionice, spajanjem sistema za učenje na daljinu, sistema za adaptivno elektronsko obrazovanje sa sistemom za praćenje aktivnosti studenata.

6. Literatura

(Evans, 2011) Dave Evans, "10 technology trends that will change the world in 10 years", Cisco Internet Business Solutions Group (IBSG), 2011.

(Madam et al. 2004) Madan A., Caneel R., Pentland A., "GroupMedia: Distributed Multi-modal Interfaces", ICMI, 2004.

(DeVaul et al. 2003) DeVaul, R. W., Pentland, S., "The MIThril Real-Time Context Engine and Activity Classification", Technical Report, MIT Media Lab, 2003

(Cooper et al. 1956) Cooper, J. B, Siegel, Helen E, "The galvanic skin response as a measure of emotion in prejudice", Journal of Psychology: Interdisciplinary and Applied, 42, 1956, 149-155, 1956.

(Healey & Picard, 1998) Healey J., R. W. Picard. "Startlecam: A cybernetic wearable camera", In ISWC, pages 42-49, 1998.

(Nakra, 2000) Nakra, T.M., "Inside the conductor's jacket: analysis, interpretation and musical synthesis of expressive gesture", PhD thesis, 2000.

(Healey, 2000) Healey J., "Wearable and automotive systems for affect recognition from physiology", 2000.

(Westeyny et al. 2006) Westeyny T., Prestiz P., Starnery, T., „ActionGSR: A Combination Galvanic Skin Response-Accelerometer for Physiological Measurements in Active Environments“, 10th IEEE International Symposium on Wearable Computers, 2006.

(Shi et al, 2002) Shi Y.C., Xie W.K., Xu G.Y., "Smart Remote Classroom: Creating a Revolutionary Real-time Interactive Distance Learning System", LNCS Vol. 2436, pp. 130-141, 2002.

(Pishva, 2007) Pishva, D., "Smart Classroom bring top-quality education around the globe", International Symposium on Applications and the Internet Workshops, pp. 40, 2007.

(Xie et al. 2001) Xie, W., Shi, Y., G. Xu and D. Xie, "Smart Classroom - an Intelligent Environment for Tele-education", IEEE Pacific Rim Conference on Multimedia, LNCS 2195, pp. 662–668, Beijing, 2001.

(Stiefelhagen et al. 2006) Stiefelhagen, R., Bernardin, K., Ekenel, H.K., McDonough J., Nickel K., Voit M., Wolfel A., "Audio-visual perception of a lecturer in a smart seminar room", Signal Processing, Vol. 86, pp. 3518–3533, 2006.

(Xin et al. 2010) Xin Li, Minghua Li, Liren Zeng, "Virtual Classrooms Supporting a Two-Way Synchronized Video and Audio Interaction", Lecture Notes in Computer Science, Vol. 6249, pp. 446-455, 2010.

(Gligorić et al. 2011) Gligorić N., Krčo S., Uzelac A., "Smart Classrooms: Survey and Development Directions", Konferencija o računarskim naukama i informacionim tehnologijama YUINFO, Kopaonik, 2011.

(Xu et al, 2009) Xu P., Han G., Li W., Wu Z., Zhou M., "Towards Intelligent Interaction in Classroom", Lecture Notes in Computer Science, 2009, Volume 5616, pp. 150-156, 2009.

(Sevindik, 2010) Sevindik T., "Future's learning environments in health education: The effects of smart classrooms on the academic achievements of the students at health college", Telematics and Informatics, Vol. 27, Issue 3, pp. 314–322, 2010.

(Luo et al. 2009) Luo Q., Zhou J., Wang F, Shen L., "Context Aware Multimodal Interaction Model in Standard Natural Classroom", Lecture Notes in Computer Science, 2009, Volume 5685/2009, 13-23

(Bouras et al. 2006) Bouras C., Giannaka E., Panagopoulos A., Tsiatsos T., "A platform for virtual collaboration spaces and educational communities: the case of EVE", *Multimedia Systems*, Vol. 11, Issue 3, pp. 290–303, 2006.

(Winera & Cooperstock, 2002) Winera L., Cooperstock J., "The Intelligent Classroom: changing teaching and learning with an evolving technological environment", *Computers & Education* 38, 253–266, 2002.

(Xiaojun et al. 2006) Xiaojun Bi, Yuanchun Shi, Xiaojie Chen, "uPen: A Smart Pen-like Device for Facilitating Interaction on Large Displays", *IEEE TableTop Conference on Horizontal Interactive Human-Computer Systems*, pp. 160-168, Australia, January 2006.

(Yokoi & Fujiyoshi, 2005) Yokoi, T., Fujiyoshi, H., "Virtual camera work for generating lecture video from high resolution images", *IEEE International Conference on Multimedia and Expo*, pp. 4, 2005.

(Dickson et al. 2008) Dickson P., Adrion W., Hanson A, "Whiteboard Content Extraction and Analysis for the Classroom Environment", *Tenth IEEE International Symposium on Multimedia*, pp. 702-707, 2008.

(Suo et al. 2007) Suo Y., Miyata N., Ishida T., Shi Y, "Open Smart Classroom: Extensible and Scalable Smart Space Using Web Service Technology", *In Proc. of 6th International Conference on Web-based Learning*, pp. 428-439, 2007.

(Di et al. 2008) Di D., Gang Z., Juhong X., "An Introduction to The Technology Of Blending-reality Smart Classroom", *International Symposium on Knowledge Acquisition and Modeling*, 2008.

(Gu & Shen, 2010) Gu H., Shen R., “An Effort to Boost Building an Attentive Service Environment for the Participants in Smart Classrooms”, Lecture Notes in Computer Science, Vol. 6483, pp. 361-367, 2010.

(Ren & Hu) Ren H., Xu G., “Human Action Recognition in Smart Classroom”, Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.

(Russell et al. 2003) Russell J., Bachorowski J., Fernandez-Dols J., “Facial and Vocal Expression of Emotion”, *Annu. Rev. Psychol.* 2003. 54:329–49

(Ambady & Rosenthal, 1992) N. Ambady and R. Rosenthal, “The Slices of Expressive Behavior as Predictor of Interpersonal Consequences: A Meta-Analysis”, *Psychological Bulletin*, 1992. Vol. 111, No.2, 256-274

(Pesarin et al. 2010) Pesarin A., Cristani M., Calanca P., Murino V., "A Generative Score Space for Statistical Dialog Characterization in Social Signalling", Lecture Notes in Computer Science, Volume 6218/2010, 630-639, 2010.

(Pentland, 2008) Pentland A., “Honest Signals: How They Shape Our World”, Massachusetts Institute of Technology, 2008.

(Pantic et al. 2009) Pantic M, Pentland A., Nijholt A, “Special Issue on Human Computing” *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions*, 3-6. 2009.

(Vinciarelli et al. 2009) Vinciarelli A., Pantic M., Bourlard H., "Social signal processing: Survey of an emerging domain", *Image and Vision Computing* 27, 1743–1759, 2009.

(Crystal, 1969) Crystal D., “Prosodic Systems and Intonation in English”, Cambridge , University Press, 1969.

(Scherer, 2003) Scherer K., "Scherer Vocal communication of emotion: a review of research paradigms", *Speech Communication* 40, 227–256. 2003.

(Scherer, 1979) Scherer K., "Personality markers in speech", Cambridge University Press, Cambridge, 1979.

(Glass et al. 2008) Glass, C., Merluzzi T., Biever J., Larsen K., "Cognitive assessment of social anxiety: development and validation of a self-statement questionnaire", *Cognitive Therapy and Research* 6, 37–55. 1982.

(Shrout & Fiske, 1981) Shrout P., Fiske D., "Nonverbal behaviors and social evaluation", *Journal of Personality* 49, pp. 115–128, 1981

(Keltner & Haidt, 1999) Keltner D., Haidt J., "Social functions of emotions at four levels of analysis", *Cognition and Emotion* 13, Vol 5, pp 505–521. 1999

(Russell & Fernandez-Dols, 1997) Russell J., Fernandez-Dols J., "The Psychology of Facial Expression", Cambridge University Press, Cambridge, 1997.

(Zellner, 1994) Zellner B., "Pauses and the temporal structure of speech", *Fundamentals of Speech Synthesis and Speech Recognition*, John Wiley & Sons, pp. 41–62. 1994

(Richmond & McCroskey, 1995) Richmond V., McCroskey J., "Nonverbal Behaviors in interpersonal relations", Allyn and Bacon, Bacon, NY, 1995.

(Psathas, 1995) Psathas G., "Conversation Analysis – The Study of Talk-in-Interaction", Sage Publications, Beverley Hills, 1995.

(Yule, 1996) Yule G., "Pragmatics", Oxford University Press, Oxford, 1996.

(Burgoon et al. 1995) Burgoon J, Stern L., Dillman L., "Interpersonal Adaptation: Dyadic Interaction Patterns", Cambridge University Press, Cambridge, 1995.

(Smith-Lovin & Brody, 1989) Smith-Lovin L., Brody C., “Interruptions in group discussions: the effects of gender and group composition“, *American Sociological Review* 54, Vol 31, pp 424–435. 1989

(Shriberg et al. 2001) Shriberg E., Stolcke A., Baron D., “Observations of overlap: findings and implications for automatic processing of multiparty conversation“, *Proceedings of Eurospeech*, pp. 1359–1362. 2001.

(Tranter & Reynolds, 2006) Tranter S., Reynolds D., “An overview of automatic speaker diarization systems“, *IEEE Transactions on Audio, Speech, and Language Processing* 14, 1557–1565. 2006.

(Ajmera et al. 2003) Ajmera J., McCowan I., Boulard H., “Speech/music segmentation using entropy and dynamism features in a HMM classification framework“, *Speech Communication* 40, 351–363. 2003.

(Lu et al. 2002) Lu L., Zhang H., Jiang H., “Content analysis for audio classification and segmentation“, *IEEE Transactions on Speech and Audio Processing* 10, 504–516. 2002.

(Gauvain et al. 1998) Gauvain J., Lamel L., Adda G., “Partitioning and transcription of broadcast news data“, *Proceedings of International Conference on Spoken Language Processing*, 1998, pp. 1335–1338.

(Rabiner & Schafer, 1978) Rabiner L., Schafer R., “*Digital Processing of Speech Signals*“, Prentice-Hall, Englewood Cliffs, NJ, 1978.

(Huang et al. 2001) Huang X., Acero A., Hon H., “*Spoken Language Processing*“, Prentice-Hall, Englewood Cliffs, NJ, 2001.

(Barras, 2004) Barras C., Zhu X., Meignier S., Gauvain J., “Improving speaker diarization“, Proceedings of the Rich Transcription Workshop, 2004.

(Oikonomopoulos et al. 2006) Oikonomopoulos I., Patras I., Pantic M., “Spatiotemporal salient points for visual recognition of human actions“, IEEE Transactions on Systems, Man, and Cybernetics – Part B, 36, pp. 710–719. 2006.

(Sminchisescu et al. 2006) Sminchisescu C., Kanaujia A., Metaxas D., “Conditional models for contextual human motion recognition“, Computer Vision and Image Understanding 104, 210–220. 2006.

(Bobick & Johnson, 2001) Bobick A., Johnson A., “Gait recognition using static activity-specific parameters“, Proceedings of Computer Vision and Pattern Recognition, pp. 423–430. 2001.

(Ito et al. 2005) Ito A., Wang X., Suzuki M., Makino S., “Smile and laughter recognition using speech processing and face recognition from conversation video“, Proceedings of the International Conference on Cyberworlds, pp. 437–444. 2005.

(Lepri et al. 2010) Lepri B., Kalimeri K., Pianesi F., “Honest Signals and Their Contribution to the Automatic Analysis of Personality Traits – A Comparative Study“, Proceeding HBU'10 Proceedings of the First international conference on Human behavior understanding, Springer-Verlag Berlin, Heidelberg, 2010.

(Chippendale, 2006) Chippendale P., “Towards Automatic Body Language Annotation“, Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, pp. 487–492, 2006.

(Kassin, 2003) Kassin S., “Psychology“, Prentice-Hall, Inc. USA, 2003.

(Jung, 1923) Jung C., “Psychologischen Typen,“, Rascher Verlag, Zurich – translation H.G. Baynes, 1923.

(Time, 1955) Time magazine, "The Old Wise Man", članak o Jungu, Feb. 14, 1955

(Sharma 1980) Sharma S., "Clothing behaviour, personality, and values: A correlational study" Psychological Studies, 25, 137–142. 1980.

(Rentfrow & Gosling, 2003) Rentfrow P., Gosling S., "The do re mi's of everyday life: The structure and personality correlates of music preference", Journal of Personality and Social Psychology, 84, 1236–1256. 2003.

(Eysenck & Eysenck, 1985) Eysenck H., Eysenck M., "Personality and Individual differences", New York: Plenum Press, 1985.

(Richmond & McCroskey, 1955) Richmond V., McCroskey J., "Nonverbal Behavior in Interpersonal Relations", Allyn & Bacon, Needham Heights, 1955.

(Gligorić et al. 2011) Gligorić N., Dejanović I., Krčo S., "Performance evaluation of compact binary XML representation for constrained devices", International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), Barcelona, 2011.

(Cosm, 2014) "Real-Time Open Data Web Service for the Internet of Things – Cosm", <https://cosm.com/>

(Google Chart Tool, 2014) "Google Chart Tool", <http://code.google.com/apis/chart/>

(PilotViewer, 2014) "PilotViewer", <http://www.microsoft.com/silverlight/pivotviewer/>

(Lockwood & Boudy, 1992) Lockwood P., Boudy J., "Experiments with a Nonlinear Nonlinear Spectral Subtractor (NSS), Hidden Markov Models and the projection, for robust speech recognition in cars", Speech Communication, Vol. 11, Nos. 2-3, pp. 215-228. 1992.

(Bakis et al. 1993) Bakis R., Nadas A., Nahamoo D., Picheny M., "Influence of background noise and microphone on the performance of the IBM TANGORA speech recognition system", Proc. IEEE. Acoustique, Speech, and Signal Processing, Vol. II, pp. 71-74. 1993.

(Gong, 1995) Gong Y., "Speech recognition in noisy environments: A survey", Speech Communication, vol 16, pp 261-291, 1995.

(Varner & Dickinson, 1967) Varner C., Dickinson G., "The Lecture, An Analysis and Review of Research", Adult Education Quarterly, vol.17, pp. 85-100, 1967.

(Yamada et al. 2001) Yamada T., Watanabe N., Asano F., Kitawaki N., "Voice activity detection using non-speech models and HMM composition", Hands-Free Speech Communication, Kyoto, Japan, pp. 131-134., 2001.

(Shriberg et al. 2001) Shriberg E., Stolcke A., Baron D., "Observations of overlap: findings and implications for automatic processing of multiparty conversation", Proceedings of Eurospeech, pp. 1359–1362. 2001.

(Merenje buke, 2014) "Uputstvo za merenje buke u životnoj sredini", <http://sr.scribd.com/doc/97835230/Uputstvo-Za-Merenje-Buke>

(The Economist 2010) The Economist, The Internet of Things, The Difference Engine: Chattering objects, 2010
http://www.economist.com/blogs/babbage/2010/08/internet_things

(Pentland, 2006) Pentland A., "Computational Model of Social Signaling", Human Dynamics Group, MIT Media Laboratory, Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)

(Loza et al. 2009) Loza A., Mihaylova L., Canagarajah N., Bull B., "Structural similarity-based object tracking in multimodality surveillance videos", *Machine Vision and Applications*, Volume 20 Issue 2, January 2009 Pages 71-83, 2009

(Sulaiman et al. 2012) Sulaiman N., Taib M., Lias S., Murat Z., Aris S., Hamid N., *Novel Methods for Stress Features Identification using EEG Signals*, *IJSSST*, Vol. 12, No. 1, 2012

(Terrence, 1999) Terrence L. *Feedforward Neural Network Methodology*, *Information Science and Statistics*, 1999, XVI, 340 p.

(Vinciarelli et al. 2009) Vinciarelli A., Pantic M., Bourlard H., "Social signal processing: Survey of an emerging domain", *Image and Vision Computing*, Vol. 27, pp. 1743–1759, 2009.

(Pentland, 2004) Pentland A., *Social Dynamics: Signals and Behavior*, MIT Media Laboratory Technical Note 579, appears *ICDL'04*, San Diego, Oct. 20-22

(Vinciarelli et al. 2009) Vinciarelli A., Salamin H., Pantic M., *Social Signal Processing: Understanding Social Interactions through Nonverbal Behavior Analysis*, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2009

(Caneel, 2005) Caneel R.. *Social signaling in decision making*. Master's thesis, Massachusetts Institute of Technology, 2005.

(Ulrici et al. 2010) Ulrici J., Hell E., Buzug, T., *Automatic detection of patient motion in cone-beam computed tomography*, *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2010, pp. 1257 – 1260

(Basu, 2002) Basu S., "Conversational Scene Analysis", PhD thesis, Dept. of Electrical Engineering and Computer science, MIT, 2002

(Andersson, 2004) Andersson T., “Audio classification and Content Description”, Master’s Thesis, Lulea University of Technology 2004.

(Lu et al. 2002) Lu L., Zhang H., Jiang H., “Content analysis for audio classification and segmentatio”, IEEE Transaction on Speech and Audio Processing, 10, pp. 504-516, 2002.

(Li et al. 2001) Li D., Sethi I., Dimitrova N., McGee T., Classification of general audio data for content-based retrieval, Pattern Recognition Letters 22, Vol 5, pp. 533-544, 2001.

(Hall et al. 2009) Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1. <http://www.cs.waikato.ac.nz/ml/weka/>

(Sphinx, 2013) Sphinx-4, A speech recognizer written entirely in the JavaTM programming language, <http://cmusphinx.sourceforge.net/sphinx4/index.html>

(Jtransforms, 2013) JTransforms, Java libraries for spectrum sound analysis, <http://sites.google.com/site/piotrwendykier/software/jtransforms>

(Dunbar & Burgoon, 2005) Dunbar N., Burgoon J., “Perceptions of power and interactional dominance in interpersonal relationships,” J. Social Personal Relationships, vol. 22, no. 2, pp. 207–233, 2005.

(Dunbar, 1998) Dunbar R., Grooming, Gossip, and the Evolution of Language. Harvard University Press, Cambridge, MA, 1998.

(Nass & Brave, 2004) Nass C., Brave S., Voice activated: How people are wired for speech and how computers will speak with us. MIT Press, Cambridge, MA, 2004.

(Werner & Keller, 1994) Werner S., Keller E., Prosodic aspects of speech. In E. Keller, editor, *Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State of the Art, and Future Challenges*, pages 23–40. John Wiley and Sons, NY, 1994.

(Handel, 1989) Handel S., *Listening: an introduction to the perception of auditory events*. MIT Press, Cambridge, MA, 1989.

(Know & Holden, 1963) Knox E., Holden J., An Instrument for Detecting Galvanic Skin Response to Acoustic Stimuli, *The Journal of Laryngology & Otology*, Vol 77, Issue 06, 1963, pp 497-500

(Fish & Geddes, 2003) Fish R., Geddes L., *Medical and Bioengineering Aspects of Electrical Injuries*, Lawyers & Judges Publishing Company, Inc, 2003

(Baenninger, 1997) Baenninger R., Yawning and its functions, *Temple University, Philadelphia, Pennsylvania Psychonomic Bulletin & Review*, 1997, pp. 198-207

(Salah et al. 2010) Salah A., Gevers T., Sebe N., Vinciarelli A., *Human Behavior Understanding: First International Workshop, Istanbul 2010*

(Zhu et al. 2002) Zhu Z., Ji Q., Fujimura K., Lee K., “Combining Kalman filtering and mean shift for real time eye tracking under active ir illumination,” in *Proc. Int’l Conf. Pattern Recognition*, Aug. 2002.

(Fielding, 2000) Fielding R., *Representational State Transfer (REST)*, Doktorska disertacija, Kalifornijski Univerzitet, 2000.
http://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf

(OpenCV, 2013) OpenCV (Open Source Computer Vision) library of programming functions for real time computer vision. <http://opencv.willowgarage.com/wiki/>

(McKay, 2012) McKay C., jMIR 2.0, McGill University, <http://jmir.sourceforge.net/>

(Weka Android, 2013) Weka Android, <https://github.com/rjmarsan/Weka-for-Android>

(JUNG, 2013) JUNG — the Java Universal Network/Graph Framework, <http://jung.sourceforge.net/>

(Hall & Llinas, 1997) Hall D., Llinas J., “An Introduction to Multisensor Data Fusion”, Proceedings of the IEEE, Vol. 85, no. 1, 1997.

(Steinert & Snell, 1999) Steinert Y, Snell L., Interactive lecturing: strategies for increasing participation in large group presentations, Medical Teacher, 1999, Vol. 21, No. 1

(Bayes Network Matlab, 2007) Bayes Network Toolbox for Matlab, 2007. <https://code.google.com/p/bnt/>

(Bayes Network Android, 2013) Bayes Network Simulator Android, <https://github.com/txus/bayes-android>

(Baesian Network Java, 2013) Bayesian Network Classifiers in Java, <http://sourceforge.net/projects/jbnc/?source>

(Morgan & Kingston) Morgan K., Kingston K, Effects of different teaching styles on the teacher behaviors that influence motivational climate and pupils' motivation in physical education, European Physical Education Review, 2005, Vol. 11 no. 3 257-285

(Sharma, 1997) Sharma D., Accounting students' learning conceptions and the influence of the learning–teaching context on approaches to learning, Accounting Education: An International Journal, 1997, Vol. 6, Issue 2, pp. 125-146

(IP Webcam, 2013) IP Webcam, Android App on Google Play, Pavel Khlebovich <https://play.google.com/store/apps/details?id=com.pas.webcam>

(Booch et al. 1999) Booch G., Rumbaugh J., Jacobson I., The Unified Modeling Language User Guide, Pearson Education, 1999

(Peker et al. 2001) Peker K., Divakaran A., Papathomas T., Automatic Measurement of Intensity of Motion Activity of Video Segments, in Proc. SPIE, Storage and Retrieval for Media Databases, 2001

(Witten et al. 1999) Witten, I., Frank, E., Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann, San Francisco, CA, 1999

(Dunbar, 2004) Dunbar, R., Grooming, gossip, and the evolution of language, Cambridge, MA: Harvard University Press. 2004

(Stoltzman, 2006) Stoltzman, W., Toward a Social Signaling Framework: Activity and Emphasis in Speech, Master Thesis, MIT, 2006. Available online: <http://hd.media.mit.edu/tech-reports/TR-608.pdf>

(Saul et al, 1995) Saul, L., Jordan, M., Boltzmann chains and hidden markov models, In Neural Information Processing Systems 7, (NIPS 7), 1995

(Wang et al. 2004) Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: From error measurement to structural similarity, IEEE Transactions on Image Processing, vol. 13, pp. 600-612, no. 1, Jan. 2004.

(Galton, 1982) Galton F., Finger Prints, Macmillan and co, New York, 1892.

(Refaeilzadeh et al. 2008) Refaeilzadeh, P., Thang, L., Liu, H.: Cross Validation, Arizona State University, 2008, http://www.cse.iitb.ac.in/~tarung/smt/papers_ppt/ency-cross-validation.pdf

(Kotsiantis, 2007) Kotsiantis, S., Supervised Machine Learning: A Review of Classification Techniques, in Proc. of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies, pp3-24, 2007

(Sundmaeker et al. 2010) Sundmaeker, H., Guillemin, P., Friess, P., Woelffle, S., Vision and Challenges for Realising the Internet of Things, 2010, http://www.theinternetofthings.eu/sites/default/files/Rob%20van%20Kranenburg/Clusterrbook%202009_0.pdf

(Manning, 2008) Manning C., Raghavan P., Schütze H., Introduction to Information Retrieval, Cambridge University Press. 2008. <http://nlp.stanford.edu/IR-book/html/htmledition/choosing-what-kind-of-classifier-to-use-1.html>

(Freund & Schapire, 1996) Freund Y., Schapire R.,. Experiments with a new boosting algorithm. Proc International Conference on Machine Learning, pp 148-156, Morgan Kaufmann, 1996

(Breiman, 2001) Breiman L., Random Forests. Machine Learning 45, Vol 1, pp. 5-32, 2001.

(Freund & Mason, 1999) Freund, Y., Mason, L.: The alternating decision tree learning algorithm. Proceeding of the Sixteenth International Conference on Machine Learning, Bled, Slovenia, pp. 124-133. 1999

(Quinlan, 1993) Quinlan R., C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, San Mateo, CA. 1993

(Platt, 1998) Platt J., Fast Training of Support Vector Machines using Sequential Minimal Optimization. Advances in Kernel Methods - Support Vector Learning, B. Schoelkopf, C. Burges, and A. Smola, eds., MIT Press, 1998

(Keerthi et al. 2001) Keerthi S., Shevade S., Bhattacharyya C., Murthy K., Improvements to Platt's SMO Algorithm for SVM Classifier Design. *Neural Computation*, 13 (3), pp 637-649, 2001.

(Aha & Kibler, 1991) Aha D., Kibler D., "Instance-based learning algorithms", *Machine Learning*, vol.6, pp. 37-66. 1991

(John & Langley, 1995) John G., Langley P., Estimating Continuous Distributions in Bayesian Classifiers. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. pp. 338-345. Morgan Kaufmann, San Mateo. 1995

Biografija

Nenad Gligorić je kao konsultant za Ericsson radio na FP7 projektima: HOBNET, EXALTED, OUTSMART i SmartSantander). Trenutno kao zaposleni u DunavNET-u radi na FP7 projektu SocioTal kao i na više M2M i Internet of Things projekata. Radio je kao recenzent za eminentne naučne časopise i konferencije: Adhoc & Sensor Wireless Networks, An International Journal of Computing and Informatics, Transactions on Emerging Telecommunications Technologies i esIoT. Rezultate istraživanja objavio je na priznatim naučnim skupovima/časopisima uključujući PerCom, DCOSS, CyberC; Transactions on Emerging Telecommunications Technologies i Journal of Ambient Intelligence and Smart Environments. Njegova polja istraživanja su M2M komunikacije, Internet inteligentnih uređaja i Sve-prisutno računarstvo.

Spisak objavljenih radova:

Radovi objavljeni u stranim časopisima:

1. Nhon Chu, Djelal Raouf, Bruno Corlay, Mohamed Ammari, Nenad Gligoric, Srdjan Krco, Nemanja Ognjanovic, Aleksandar Obradovic, OMA DM v1.x compliant lightweight device management for constrained M2M devices, Transactions on Emerging Telecommunications Technologies, Vol 24 Issue 5, 2013. IF 1.049 (2012)
2. Nenad Gligorić, Ana Uzelac, Srđan Krčo, Ivana Kovačević, Ana Nikodijević, Smart Classroom System for Detecting Level of Interest a Lecture Creates in a Classroom, Journal of Ambient Intelligence and Smart Environments, in press, 2014. IF 1.298 (2012)

Radovi saopšteni na skupovima međunarodnog značaja štampani u celini:

1. Nenad Gligorić, Ana Uzelac, Srđan Krčo, "Smart Classroom: Real-Time Feedback on Lecture Quality", IEEE Conference on Pervasive Computing and Communications, PerCom 2012, WiP session
2. Nenad Gligorić, Igor Dejanović, Srđan Krčo, "Performance evaluation of compact

binary XML representation for constrained devices”, International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), Barcelona, 2011.

3. Miljan Vučetić, Ana Uzelac, Nenad Gligorić, “E-health Transformation Model in Serbia: Design, Architecture and Developing”, International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC 2011), Beijing, 2011.

4. Nenad Gligorić, Srđan Krčo, Dejan Drajić, Stevan Jokić, Bojana Jakovljević, “M2M Device Management in LTE Networks”, XIX Telecommunications Forum TELFOR, Beograd, 2011.

5. Stevan Jokić, Srđan Krčo, Jelena Vučković, Nenad Gligorić, Dejan Drajić, “Evaluation of an XML Database Based Resource Directory Performance“, XIX Telecommunications Forum TELFOR, Beograd, 2011.

6. Tomislav Dimčić, Srđan Krčo, Nenad Gligorić, “CoAP (Constrained Application Protocol) implementation in M2M Environmental Monitoring System”, 2nd International Conference on Information Society Technology, 2012

7. Nenad Gligorić, Tomislav Dimčić, Dejan Drajić, Srđan Krčo, Nhon Chu, “Application-Layer Security Mechanism for M2M communication over SMS”, XX Telecommunications Forum TELFOR, Beograd, 2012.

8. Nenad Gligorić, Tomislav Dimčić, Dejan Drajić, Srđan Krčo, Igor Dejanović, Nhon Chu, Aleksandar Obradović, “CoAP Over SMS: Performance Evaluation for Machine to Machine Communication”, XX Telecommunications Forum TELFOR, Beograd, 2012.

9. Nenad Gligorić, Srđan Krčo, Ignacio EliceGUI, Carmen López, Luis Sánchez, Michele Nati, Rob van Kranenburg, M. Victoria Moreno, Davide, “SocIoTal: Creating a Citizen-Centric Internet of Things”, 4th International Conference on Information Society Technology, 2014

Radovi objavljeni u domaćim časopisima:

1. Stevan Jokić, Srđan Krčo, Igor Dejanović, Jelena Vučković, Nenad Gligorić, and Dejan Drajić, “Evaluation of a Document Oriented Resource Directory Performance”, Telfor Journal, Vol. 4, No. 2, 2012

2. Nenad Gligorić, Ana Uzelac, Miljan Vučetić, Ljiljana Milosavljević, “Advance Driver Assistance Systems Based on Interaction with Other Vehicles and Infrastructure Communication”, INFO M, Vol. 9, br. 34, str. 18-22, Beograd, 2010.

3. Miljan Vučetić, Ana Uzelac, Nenad Gligorić, Vladimir Laketa, “Transformation E-Health model in Serbia”, INFO M, Vol. 9, br. 36, str. 28-31, Beograd, 2010.

4. Nenad Gligorić, Ana Uzelac, Dejan Drajić, Sanja Vuković, “Application of Wireless

Technology in Traffic”, Singidunum revija, ISSN: 1820-8819, Vol. 7, pp. 169-179, Beograd, 2010.

5. Nenad Gligorić, Tomislav Dimčić, Srđan Krčo, Vladislav Dimčić, Jelena Vasković, Iva Vojinović, “Internet of Things enabled LED lamp controlled by satisfaction of students in a classroom”, The IPSI BgD Transactions on Internet Research, 2014

Radovi saopšteni na skupu nacionalnog značaja štampani u celini:

1. Nenad Gligorić, Srđan Krčo, Ana Uzelac, “Model za praćenje aktivnosti u učionici primenom tehnologije Interneta Inteligentnih Uređaja”, Konferencija o računarskim naukama i informacionim tehnologijama YUINFO, Kopaonik, 2012.

2. Nenad Gligorić, Srđan Krčo, Ana Uzelac, “Smart Classrooms: Survey and Development Directions”, Konferencija o računarskim naukama i informacionim tehnologijama YUINFO, Kopaonik, 2011.

3. Dragan Zoranović, Ana Uzelac, Nenad Gligorić, “Aplikacija za podršku paralelno sinhronizovane reprodukcije medijskih izvora”, XXXVIII Simpozijum o operacionim istraživanjima, SYM-OP-IS 2011, Zbornik radova, str. 302-305, Zlatibor 2011.

4. Ana Uzelac, Dragan Zoranović, Nenad Gligorić, Miljan Vučetić, Sanja Vuković, “Unapređenje zdravstvenog sistema zemalja u razvoju primenom mobilnih tehnologija”, Archives for Technical Sciences, Zbornik radova, vol 5. 2011.

Prilog 1.

Izjava o autorstvu

Potpisan Nenad R. Gligorić, broj indeksa 03/08

Izjavljujem

da je doktorska disertacija pod naslovom

Praćenje aktivnosti studenata primenom Interneta inteligentnih uređaja

- rezultat sopstvenog istraživačkog rada,
- da predložena disertacija u celini ni u delovima nije bila predložena za dobijanje bilo koje diplome prema studijskim programima drugih visokoškolskih ustanova,
- da su rezultati korektno navedeni i
- da nisam kršila autorska prava i koristila intelektualnu svojinu drugih lica.

U Beogradu, 08.04.2014.

Potpis doktoranta



Prilog 2.

Izjava o istovetnosti štampane i elektronske verzije doktorskog rada

Ime i prezime autora: Nenad R. Gligorić

Broj indeksa: 03/08

Studijski program: Elektronsko poslovanje

Naslov rada: Praćenje aktivnosti studenata primenom Interneta inteligentnih uređaja

Mentor: Prof. dr Srđan Krčo

Potpisan Nenad Gligorić

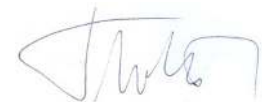
Izjavljujem da je štampana verzija mog doktorskog rada istovetna elektronskoj verziji koju sam predao za objavljivanje na portalu **Digitalnog repozitorijuma Univerziteta u Beogradu**.

Dozvoljavam da se objave moji lični podaci vezani za dobijanje akademskog zvanja doktora nauka, kao što su ime i prezime, godina i mesto rođenja i datum odbrane rada.

Ovi lični podaci mogu se objaviti na mrežnim stranicama digitalne biblioteke, u elektronskom katalogu i u publikacijama Univerziteta u Beogradu.

U Beogradu, 08.04.2014.

Potpis doktoranta



Prilog 3.

Izjava o korišćenju

Ovlašćujem Univerzitetsku biblioteku „Svetozar Marković“ da u Digitalni repozitorijum Univerziteta u Beogradu unese moju doktorsku disertaciju pod naslovom:

Praćenje aktivnosti studenata primenom Interneta inteligentnih uređaja

koja je moje autorsko delo.

Disertaciju sa svim priložima predao sam u elektronskom formatu pogodnom za trajno arhiviranje.

Moju doktorsku disertaciju pohranjenu u Digitalni repozitorijum Univerziteta u Beogradu mogu da koriste svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (Creative Commons) za koju sam se odlučila.

1. Autorstvo
2. Autorstvo - nekomercijalno
3. Autorstvo – nekomercijalno – bez prerade
4. Autorstvo – nekomercijalno – deliti pod istim uslovima
5. Autorstvo – bez prerade
6. Autorstvo – deliti pod istim uslovima

U Beogradu, 08.04.2014.

Potpis doktoranta

