



### FIȘA DISCIPLINEI

#### 1. Date despre program

1.1 Instituția de învățământ superior	Universitatea Națională de Știință și Tehnologie Politehnica București
1.2 Facultatea	Electronică, Telecomunicații și Tehnologia Informației
1.3 Departamentul	Dispozitive, Circuite și Arhitecturi Electronice
1.4 Domeniul de studii	Inginerie Electronică, Telecomunicații și Tehnologii Informaționale
1.5 Ciclul de studii	Masterat
1.6 Specializarea	Tehnologii Multimedia în Aplicații de Biometrie și Securitatea Informației

#### 2. Date despre disciplină

2.1 Denumirea disciplinei (ro)				Interfețe de comunicare prin voce cu sisteme inteligente			
(en)				Voice Communication Interfaces with Intelligent Systems			
2.2 Titularul activităților de curs				Prof. dr. ing. Dragoș BURILEANU, Conf. dr. ing. Horia CUCU			
2.3 Titularul activităților de seminar / laborator				Conf. dr. ing. Horia CUCU			
2.4 Anul de studiu	1	2.5 Semestrul	II	2.6. Tipul de evaluare	E	2.7 Regimul disciplinei	Ob
2.8 Tipul disciplinei	DA	2.9 Codul disciplinei	UPB.04.M2.O.20-03	2.10 Tipul de notare	Nota		

#### 3. Timpul total estimat (ore pe semestru al activităților didactice)

3.1 Număr de ore pe săptămână	3	Din care: 3.2 curs	2.00	3.3 seminar/laborator	1
3.4 Total ore din planul de învățământ	42.00	Din care: 3.5 curs	28	3.6 seminar/laborator	14
Distribuția fondului de timp:					ore
Studiul după manual, suport de curs, bibliografie și notițe					79
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate					
Pregătire seminarii/ laboratoare/proiecte, teme, referate, portofolii și eseuri					
Tutorat					0
Examinări					4
Alte activități (dacă există):					0
3.7 Total ore studiu individual	83.00				
3.8 Total ore pe semestru	125				
3.9 Numărul de credite	5				

#### 4. Precondiții (acolo unde este cazul)

4.1 de curriculum	Parcurgerea următoarelor discipline: – Tehnologii biometrice. Analiza semnalului vocal și a semnalelor biologice – Inteligență artificială I: sisteme clasice de învățare automată
-------------------	--



4.2 de rezultate ale învățării	Acumularea următoarelor cunoștințe generale: – cunoștințe de bază despre analiza și prelucrarea semnalului vocal; – concepte fundamentale despre modele și tehnici de învățare automată.
--------------------------------	--

**5. Condiții necesare pentru desfășurarea optimă a activităților didactice** (acolo unde este cazul)

5.1 Curs	– Cursul se va desfășura într-o sală dotată cu videoproiector și computer.
5.2 Seminar/ Laborator/Proiect	– Laboratorul se va desfășura într-o sală cu dotare specifică, care trebuie să includă: videoproiector, computer, software și dispozitive pentru achiziția și prelucrarea de semnal vocal. – Prezența obligatorie la laboratoare (conform regulamentului studiilor universitare de masterat în UNSTPB).

**6. Obiectiv general** (*Se referă la intențiile profesorilor pentru studenți, la ceea ce studenții vor fi învățați în timpul cursului. Oferă o orientare cu privire la locul cursului în cadrul domeniului științific abordat, precum și la rolul pe care acesta îl are în cadrul specializării studiate. Vor fi descrise de o manieră generală tematicile abordate, justificarea includerii cursului în planul de învățământ al specializării studiate etc.*)

Prima parte a cursului își propune să ofere studenților noțiuni de bază legate de principiile și paradigmele interacțiunii om-mașină, precum și cunoștințe teoretice și practice despre interfețele de comunicare prin voce cu diverse sisteme inteligente. Sunt introduse cunoștințe legate de prelucrarea limbajului natural, sinteza automată a vorbirii și dezvoltarea și evaluarea sistemelor evoluate de dialog și se discută conceptul de multimodalitate în realizarea interfețelor interactive actuale.

Partea a doua a cursului are ca scop familiarizarea studenților cu recunoașterea automată a vorbirii (procesul prin care vorbirea este transformată în text), ca element de bază într-o interfață de comunicare prin voce între om și un sistem inteligent. Cursul prezintă arhitecturi de rețele neurale profunde utilizate cu succes pentru extragerea de caracteristici din vorbire și pentru recunoașterea propriu-zisă. Sunt descrise, de asemenea, resursele necesare învățării și adaptării unei soluții de recunoaștere / transcriere de vorbire, ca și modalitățile de post-procesare a textului rezultat.

Laboratorul prezintă modalitatea de implementare a unui sistem de tip chatbot vocal. Implementarea se realizează în limbajul de programare Python și utilizează ca punct de pornire soluții de transcriere de vorbire și sinteză automată deja existente.

**7. Competențe** (*Capacitatea dovedită de a utiliza cunoștințe, aptitudini și abilități personale, sociale și/sau metodologice în situații de muncă sau de studiu și pentru dezvoltarea profesională și personală. Reflectă cerințele angajatorilor.*)



<b>Specifice</b>	<ul style="list-style-type: none"><li>– Demonstrează că deține cunoștințe de bază privind prelucrarea semnalelor și învățarea automată aplicate la realizarea interfețelor de comunicare prin voce cu diverse sisteme inteligente.</li><li>– Aplică în practică cunoștințele teoretice dobândite și utilizează programe specializate pentru prelucrarea semnalului vocal și dezvoltarea de interfețe de comunicare prin voce.</li><li>– Aplică metode și instrumente standardizate, specifice domeniilor prelucrării semnelor vocale și a limbajului natural, pentru realizarea procesului de evaluare a unei situații, în funcție de problemele de rezolvat și identifică soluții.</li><li>– Argumentează și analizează coerent și corect contextul de aplicare a cunoștințelor de bază ale domeniului interfețelor om-mașină, utilizând concepte cheie ale disciplinei și metodologia specifică.</li><li>– Comunicare orală și în scris în limba română: utilizează vocabularul științific specific domeniului studiat, în vederea comunicării eficiente și corecte, în scris și oral.</li><li>– Comunicare orală și în scris într-o limbă străină (engleză): demonstrează înțelegerea și aplicarea corectă a vocabularului aferent domeniului studiat, într-o limbă străină.</li></ul>
<b>Transversale (generale)</b>	<ul style="list-style-type: none"><li>– Comunică eficient, în special în timpul orelor de aplicații, coordonându-și eforturile cu ceilalți pentru rezolvarea de situații problemă de complexitate medie.</li><li>– Autonomie și gândire critică: abilitatea de a gândi în termeni științifici, de a căuta și analiza date în mod independent, de a identifica soluții, precum și de a desprinde și prezenta concluzii.</li><li>– Capacitate de analiză și sinteză: prezintă în mod sintetic cunoștințele dobândite, ca urmare a unui proces de analiză sistematică.</li><li>– Respectă principiile de etică academică: în activitatea de documentare citează corect sursele bibliografice utilizate.</li><li>– Pune în practică elemente de inteligență emoțională în gestionarea socio-emoțională adecvată a unor situații din viața academică, demonstrând stăpânire de sine și obiectivitate în luarea deciziilor sau în situații de stres.</li></ul>

**8. Rezultatele învățării** (Sunt enunțuri sintetice referitoare la ceea ce un student va fi capabil să facă sau să demonstreze la finalizarea unui curs. Rezultatele învățării reflectă realizările studentului și mai puțin intențiile profesorului. Rezultatele învățării informează studenții despre ceea ce se așteaptă de la ei din punct de vedere al performanței, pentru a obține notele și creditele dorite. Sunt definite în termeni concreți, folosind verbe similare exemplurilor de mai jos și indică ceea ce se va urmări prin evaluare. Rezultatele învățării vor fi astfel redactate încât să fie evidențiată clar relația față de competențele definite la punctul 7.)



<b>Cunoștințe</b>	<p><i>Rezultatul asimilării de informații prin învățare. Cunoștințele reprezintă ansamblul de fapte, principii, teorii și practici legate de un anumit domeniu de muncă sau de studiu. Pot fi teoretice și/sau faptice.</i></p> <ul style="list-style-type: none"><li>– Definește corect noțiunile de bază ale domeniului interfețelor de comunicare prin voce cu sisteme inteligente: principiile interacțiunii prin voce om-mașină, prelucrarea limbajului natural, sinteza automată a vorbirii, recunoașterea automată a vorbirii etc.</li><li>– Definește și utilizează elementele de bază legate de producerea și percepția vorbirii.</li><li>– Înțelege particularitățile sistemelor bazate pe tehnologia vorbirii și a metodelor de dezvoltare și implementare a acestor sisteme.</li><li>– Este capabil să utilizeze corect principalele modalități de reprezentare parametrică a semnalului vocal și de extragere a caracteristicilor acestuia.</li><li>– Înțelege atât conceptele de bază ale domeniului învățării automate, cât și tehnicile moderne bazate pe rețele neuronale profunde, aplicate în realizarea interfețelor prin voce.</li><li>– Descrie în mod corespunzător conceptele fundamentale legate de sistemele inteligente de dialog și sursele de eroare ce pot apărea în funcționarea acestora.</li><li>– Evidențiază modalitățile de testare și evaluare pentru o interfață de comunicare prin voce.</li></ul>
<b>Aptitudini</b>	<p><i>Capacitatea de a aplica cunoștințe și de a utiliza know-how pentru a duce la îndeplinire sarcini și a rezolva probleme. Aptitudinile sunt descrise ca fiind cognitive (implicând utilizarea gândirii logice, intuitive și creative) sau practice (implicând dexteritate manuală și utilizarea de metode, materiale, unelte și instrumente).</i></p> <ul style="list-style-type: none"><li>– Selectează și grupează informații relevante într-un context dat, putând să descrie corespunzător diverse aspecte teoretice sau practice ale domeniului interfețelor de comunicare prin voce.</li><li>– Utilizează argumentat conceptele specifice domeniului, în vederea abordării corecte a unor probleme.</li><li>– Verifică experimental soluțiile identificate pentru dezvoltarea de module software și interfețe dedicate pentru sisteme de analiză și prelucrare a semnalului vocal.</li><li>– Formulează concluzii corecte asupra rezultatelor experimentele realizate.</li><li>– Argumentează modul de rezolvare și soluțiile utilizate pentru rezolvarea unor probleme.</li></ul>
<b>Responsabilitate și autonomie</b>	<p><i>Capacitatea cursantului de a aplica în mod autonom și responsabil cunoștințele și aptitudinile sale.</i></p> <ul style="list-style-type: none"><li>– Selectează surse bibliografice potrivite și le analizează.</li><li>– Respectă principiile de etică academică, citând corect sursele bibliografice utilizate.</li><li>– Demonstrează receptivitate pentru contexte noi de învățare.</li><li>– Manifestă colaborare cu ceilalți colegi și cadre didactice în desfășurarea activităților didactice.</li><li>– Demonstrează autonomie în organizarea contextului de învățare și a problemelor de rezolvat.</li><li>– Conștientizează valoarea contribuției sale în domeniul ingineriei la identificarea de soluții viabile care să rezolve probleme din viața socială și economică.</li><li>– Analizează oportunități de afaceri sau de dezvoltare antreprenorială, pornind de la cunoștințele dobândite în domeniul interfețelor de comunicare prin voce cu sisteme inteligente.</li><li>– Demonstrează abilități de management ale situațiilor din viața reală (de exemplu gestionarea corectă a timpului de învățare).</li></ul>

**9. Metode de predare** (Se vor avea în vedere metode care să asigure predarea centrată pe student. Se va descrie modul în care se asigură participarea studenților la stabilirea propriului parcurs de învățare, cum se identifică eventualele rămăneri în urmă și ce măsuri remediale se adoptă în astfel de cazuri.)



– Cursurile sunt predate într-o manieră interactivă, fiind încurajată participarea activă a studenților. Sunt folosite atât metode clasice de predare (prelegerea și expunerea), utilizând prezentări PowerPoint prin intermediul mijloacelor multimedia, cât și interactive, bazate pe întrebări – răspunsuri și feedback-ul studenților, adaptând permanent demersul pedagogic la posibilitățile de asimilare și învățare a studenților (prin repetarea suplimentară a anumitor noțiuni și concepte, dacă acest lucru se dovedește necesar). Fiecare curs debutează cu recapitularea succintă a capitolelor anterioare, cu accent asupra noțiunilor parcurse la ultimul curs. Prezentările utilizează numeroase imagini și scheme, astfel încât informațiile prezentate să fie cât mai ușor de înțeles și asimilat. Materialele complete de curs sunt disponibile în format electronic pe platforma Moodle a facultății.

– Predarea cunoștințelor în cadrul orelor de laborator se bazează pe comunicarea orală și explicarea detaliată a metodelor utilizate și a rezultatelor obținute, într-o manieră permanent interactivă. Studenții implementează și evaluează independent problemele de studiu prin utilizarea calculatorului, a mediului software și a echipamentelor dedicate (atunci când este cazul). Aplicațiile realizate îi ajută pe studenți în dezvoltarea unor relații optime de comunicare într-un climat favorabil învățării prin descoperire. Materialele de laborator sunt disponibile studenților sub formă electronică pe platforma Moodle a facultății.

## 10. Conținuturi

CURS		
Capitolul	Conținutul	Nr. ore
1	“Introducere. Interfețe om-mașină”	2
2	“Interfețe de comunicare prin voce” – Generalități, istoric. Interfețe cu arhitectură centrată pe rețea. Interfețe de tip “embedded”. Asistenți personali inteligenți	2
3	“Prelucrarea limbajului natural (NLP)” – Definiții. Aplicații. Tehnici folosite în prelucrarea limbajului natural	2
4	“Sinteza vorbirii pentru interfețele de comunicare prin voce” – Sinteza automată a vorbirii. Sinteza vorbirii pornind de la text: sisteme clasice și sisteme bazate pe rețele neurale profunde	5
5	“Sisteme de dialog” – Principii de bază ale interacțiunii prin voce om-calculator. Strategii de dialog. Gestionarea erorilor. Interfețe multimodale	3
6	“Introducere în recunoașterea automată a vorbirii (RAV)” – Sarcina de recunoaștere a vorbirii. Aplicații. Istoric al modelelor	2
7	“Seturi de date și metrici de evaluare pentru RAV”	2
8	“Sisteme clasice de RAV: modele pipeline”	4
9	“Sisteme de RAV end-to-end: clasificare temporală conexionistă (CTC)”	2
10	“Sisteme de RAV end-to-end: modele de tip codor-decodor”	2
11	“Antrenarea nesupervizată a sistemelor de RAV”	2
		<b>Total:</b> 28



**Bibliografie:**

1. D. Burileanu, H. Cucu, *Interfețe de comunicare prin voce cu sisteme inteligente*, suport de curs electronic pe platforma Moodle a facultății de ETTI: <https://curs.upb.ro/>
2. D. Burileanu, “Spoken Language Interfaces for Embedded Applications”, *Human Factors and Voice Interactive Systems* (D. Gardner-Bonneau and H. Blanchard – Eds.), 2nd Edition, Springer US, New York, pp. 135-162, 2008.
3. D. Burileanu, C. Negrescu, M. Surmei, “Recent Advances in Romanian Language Text-to-Speech Synthesis”, *Proceedings of the Romanian Academy, Series A*, Editura Academiei Române, București, Vol. 11, No. 1, pp. 92-99, 2010.
4. C. Pearl, *Designing Voice User Interfaces: Principles of Conversational Experiences*, O’Reilly Media, 2016.
5. Al.L. Georgescu, H. Cucu, C. Burileanu, “Kaldi-based DNN Architectures for Speech Recognition in Romanian”, *Proc. of the 10th Int. Conf. on Speech Technology and Human-Computer Dialogue (SpeD)*, Timișoara, pp. 1-6, Oct. 10-12, 2019, IEEE NY.
6. W. Chan, N. Jaitly, Q. Le, O. Vinyals, “Listen, Attend and Spell: A Neural Network for Large Vocabulary Conversational Speech Recognition”, *Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, pp. 4960-4964, 2016, IEEE NY.
7. S. Schneider, A. Baevski, R. Collobert, M. Auli, “wav2vec: Unsupervised Pre-training for Speech Recognition”, arXiv preprint, arXiv:1904.05862, 2019.
8. A. Hannun et al., “Deep Speech: Scaling up End-to-End Speech Recognition”, arXiv preprint, arXiv:1412.5567, 2014.
9. A. van den Oord, Y. Li, O. Vinyals. “Representation Learning with Contrastive Predictive Coding”, arXiv preprint, arXiv:1807.03748, 2018.

**LABORATOR**

Nr. crt.	Conținutul	Nr. ore
1	Protecția muncii. Introducere despre chatboturi. Principii legate de configurarea unor astfel de sisteme; configurarea mediului de dezvoltare	2
2	Proiectarea aplicației de tip chatbot vocal (dataflow, structură aplicație, definirea scenariilor)	2
3	Implementare logică NLP	2
4	Integrare cu TTS (text-to-speech) API: gestionarea dialogului, utilizarea dinamică a diferitelor funcționalități	2
5	Integrare cu STT (speech-to-text) API: lucrul cu diferite surse audio, aplicații diverse	2
6	Implementarea aplicației finale: îmbinarea componentelor (claselor / funcțiilor) din laboratoarele anterioare	2
7	Colocviu final	2
	<b>Total:</b>	14



### Bibliografie:

1. H. Cucu, *Interfețe de comunicare prin voce cu sisteme inteligente – Platforme de laborator*, disponibile în format electronic pe platforma Moodle a facultății de ETTI: <https://curs.upb.ro/>
2. Al.L. Georgescu, A. Pappalardo, H. Cucu, M. Blott, “Performance vs. Hardware Requirements in State-of-the-Art Automatic Speech Recognition”, *EURASIP Journal on Audio, Speech, and Music Processing*, 2021(1), 1-30.
3. Google Cloud STT: <https://cloud.google.com/speech-to-text/docs>
4. Google Cloud TTS: <https://cloud.google.com/text-to-speech/docs>
5. Google Cloud NLP: <https://cloud.google.com/natural-language?hl=en>
6. SpaCy NLP module: <https://spacy.io/models/ro>
7. SSML standard: <https://www.w3.org/TR/speech-synthesis11/>
8. Asyncio python library: <https://docs.python.org/3/library/asyncio.html>
9. Audio in python: <https://wiki.python.org/moin/Audio/>
10. Zevo Live STT API: <https://bit.ly/47Ywu4b>
11. Zevo Live TTS API: <https://bit.ly/3RqWgsj>
12. Docker intro: <https://docs.docker.com/get-started/overview/>

### 11. Evaluare

Tip activitate	11.1 Criterii de evaluare	11.2 Metode de evaluare	11.3 Pondere din nota finală
11.4 Curs	Cunoașterea noțiunilor teoretice fundamentale legate de interfețele de comunicare prin voce cu sisteme inteligente. Cunoașterea modului de aplicare a teoriei la rezolvarea unor probleme specifice domeniului.	Examen scris în sesiunea de examene.	50%
	Cunoașterea noțiunilor teoretice fundamentale legate de interfețele om-mașină. Capacitatea de a realiza o sinteză relevantă din punct de vedere științific, privind o interfață particulară.	Temă de casă individuală (descrierea unui tip de interfață om-mașină).	20%
11.5 Seminar/laborator/proiect	Înțelegerea tehnicilor fundamentale de implementare a unei interfețe de comunicare prin voce. Cunoașterea modului de simulare și de implementare practică (pe calculator) a metodelor studiate, cu ajutorul unor medii de programare evaluate.	Colocviu final de laborator (test pe calculator).	30%
11.6 Condiții de promovare			
– Obținerea a 50% din punctajul total. – Realizarea obligațiilor caracteristice activității de laborator (participarea la lucrările planificate).			

**12. Coroborarea conținutului disciplinei cu așteptările reprezentanților angajatorilor și asociațiilor profesionale reprezentative din domeniul aferent programului, precum și cu stadiul actual al cunoașterii în domeniul științific abordat și practicile în instituții de învățământ superior din Spațiul European al Învățământului Superior (SEIS)**



**Universitatea Națională de Știință și Tehnologie Politehnica București**

**Facultatea de Electronică, Telecomunicații și  
Tehnologia Informației**



Programa disciplinei oferă studenților suportul teoretic și practic necesar atât cunoașterii principiilor și tehnicilor de bază în prelucrarea semnalului vocal și a limbajului natural, cât și proiectării și evaluării unor interfețe de comunicare prin voce cu sisteme inteligente, bazate pe recunoașterea și sinteza automată a vorbirii. Disciplina răspunde concret cerințelor actuale de dezvoltare și evoluție a economiei europene a serviciilor din domeniul ICT, dar și practicilor curente din cadrul celor mai avansate instituții de învățământ superior din Europa.

În contextul progresului actual al tehnologiei informației și Inteligenței Artificiale, domeniile de activitate vizate sunt extrem de numeroase, de la aplicații de “consum” (tehnologii de vorbire pentru înglobarea în diverse aparate electrocasnice sau terminale mobile inteligente), domeniul medical (produse și tehnologii de analiza și prelucrarea semnalelor, comanda roboților medicali, sisteme de dictare a prescripțiilor medicale, indexarea bazelor de date medicale), domeniul de indexare audio pentru aplicații multimedia, până la roboți și asistenți virtuali inteligenți ce includ interfețe interactive bazate pe recunoașterea vorbirii și răspuns prin voce.

Se asigură astfel absolvenților competențe adecvate cu necesitățile calificărilor actuale și o pregătire științifică și tehnică modernă, de calitate și competitivă, care să le permită angajarea rapidă după absolvire, disciplina fiind perfect încadrată în politica Universității Naționale de Știință și Tehnologie POLITEHNICA București, atât din punctul de vedere al conținutului și structurii, cât și din punctul de vedere al aptitudinilor și deschiderii internaționale oferite studenților. Posibili angajatori vizează atât mediul academic (profil didactic și de cercetare), cât și mediul de cercetare-dezvoltare din instituțiile de stat și private care utilizează tehnologiile vorbirii (în mod special recunoașterea și sinteza automată), în realizarea unor interfețe interactive sau a unor sisteme evaluate de dialog.

Data completării

Titular de curs

Titular(i) de aplicații

Prof. dr. ing. Dragoș BURILEANU

Conf. dr. ing. Horia CUCU

Data avizării în departament

Director de departament

31.10.2024

Prof. Dr. Claudius DAN

Data aprobării în Consiliul Facultății

Decan

01.11.2024

Prof. Dr. Mihnea Udrea