



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)				Proiect de cercetare-dezvoltare în tehnologia vorbirii			
(en)				Research Project in Speech Technology			
2.2 Titularul activităților de curs				–			
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	V	2.7 Regimul disciplinei	Ob
2.8 Tipul disciplinei		DS	2.9 Codul disciplinei	UPB.04.M2.O.20-10	2.10 Tipul de notare	Nota	

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

3.1 Număr de ore pe săptămână	2	Din care: 3.2 curs	0.00	3.3 seminar/laborator	2
3.4 Total ore din planul de învățământ	28.00	Din care: 3.5 curs	0	3.6 seminar/laborator	28
Distribuția fondului de timp:					ore
Studiul după manual, suport de curs, bibliografie și notițe					
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate					41
Pregătire seminarii/ laboratoare/proiecte, teme, referate, portofolii și eseuri					
Tutorat					2
Examinări					4
Alte activități (dacă există):					0
3.7 Total ore studiu individual	47.00				
3.8 Total ore pe semestru	75				
3.9 Numărul de credite	3				

4. Precondiții (acolo unde este cazul)

4.1 de curriculum	Parcursarea 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: <ul style="list-style-type: none">– programare în limbaj Python;– 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	Nu este cazul.
5.2 Seminar/ Laborator/Proiect	– Proiectul 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.

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.*)

Proiectul de cercetare are ca scop familiarizarea studenților cu conceptul de recunoaștere a vorbitorului, într-o manieră foarte practică: realizarea individual sau în echipe de doi studenți a unui sistem personalizat de recunoașterea vorbitorului. Sunt prezentate diversele scenarii de lucru (de exemplu, identificare vs. verificare, set închis vs. set deschis de vorbitori) și este abordat unul dintre acestea. Studenții utilizează seturi de date disponibile public, atât cu vorbire în limba română, cât și multilingve și arhitecturi de rețele neurale moderne. Studenții au și opțiunea de a-și crea propriul set de date de vorbire înregistrată astfel încât sistemul realizat să poată recunoaște și vocea lor. Proiectul asigură pregătirea studenților în vederea formulării ipotezelor experimentale, proiectării experimentelor aferente, realizării practice a experimentelor și raportării rezultatelor științifice.

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.*)

Specifice	<ul style="list-style-type: none">– Demonstrează că deține cunoștințe de bază privind prelucrarea semnalului vocal și învățarea automată aplicate la realizarea sistemelor de identificare sau verificare a vorbitorului.– Aplică în practică cunoștințele teoretice dobândite și utilizează programe specializate pentru realizarea de experimente științifice de identificare sau verificare a vorbitorului.– Aplică metode și instrumente standardizate, specifice domeniilor prelucrării semnalului vocal și a limbajului natural, pentru realizarea procesului de evaluare a unei situații, în funcție de problemele de rezolvat și identifică soluții.– Argumentează și analizează coerent și corect contextul de aplicare a cunoștințelor de bază ale domeniilor prelucrării semnalului vocal și învățării automate, utilizând concepte cheie ale disciplinei și metodologia specifică.– Comunicare orală și în scris în limba română: utilizează vocabularul științific specific domeniului studiat, în vederea comunicării eficiente și corecte, în scris și oral.– Comunicare orală și în scris într-o limbă străină (engleză): demonstrează înțelegerea și aplicarea corectă a vocabularului aferent domeniului studiat, într-o limbă străină.
------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Transversale (generale)	<ul style="list-style-type: none">– Comunică eficient în timpul orelor de aplicații coordonându-și eforturile cu ceilalți pentru rezolvarea de situații problemă de complexitate medie.– 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.– Capacitate de analiză și sinteză: prezintă în mod sintetic cunoștințele dobândite, ca urmare a unui proces de analiză sistematică.– Respectă principiile de etică academică: în activitatea de documentare citează corect sursele bibliografice utilizate.– 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.
--------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

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.)

Cunoștințe	<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">– Definește corect noțiunile de bază aferente domeniului recunoașterii vorbitorului: parametrii de semnal vocal, tipul de recunoaștere (identificare vs. verificare), setul de vorbitori (închis vs. deschis), metricile de evaluare etc.– Definește și utilizează corect, în contextul recunoașterii vorbitorului, elementele de bază legate de prelucrarea semnalului vocal și învățarea automată.– Înțelege particularitățile sistemelor de recunoașterea vorbitorului și a metodelor de proiectare și implementare a acestor sisteme.– Înțelege atât conceptele de bază ale domeniului învățării automate, cât și tehnicile moderne bazate pe rețele neurale profunde, aplicate în realizarea sistemelor de recunoașterea vorbitorului.– Descrie coerent și corect modul de funcționare a sistemelor de recunoașterea vorbitorului și sursele de eroare ce pot apărea în funcționarea acestora.– Evidențiază modalitățile de evaluare și metricile de performanță pentru sistemele de recunoașterea vorbitorului.
Aptitudini	<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">– Selectează și grupează informații relevante într-un context dat, putând să descrie corespunzător diverse aspecte teoretice sau practice ale sistemelor de recunoaștere a vorbitorului.– Utilizează argumentat conceptele specifice domeniului, în vederea abordării corecte a unor probleme.– Verifică experimental soluțiile identificate pentru problemele apărute în timpul dezvoltării sistemelor de recunoaștere a vorbitorului.– Formulează concluzii corecte asupra rezultatelor experimentele realizate.– Argumentează modul de rezolvare și soluțiile utilizate pentru rezolvarea unor probleme.



Responsabilitate și autonomie	<i>Capacitatea cursantului de a aplica în mod autonom și responsabil cunoștințele și aptitudinile sale.</i>
	– Selectează surse bibliografice potrivite și le analizează.
	– Respectă principiile de etică academică, citând corect sursele bibliografice utilizate.
	– Demonstrează receptivitate pentru contexte noi de învățare.
	– Manifestă colaborare cu ceilalți colegi și cadre didactice în desfășurarea activităților didactice.
	– Demonstrează autonomie în organizarea contextului de învățare și a problemelor de rezolvat.
	– Conștientizează valoarea contribuției sale în domeniul ingineriei la identificarea de soluții viabile care să rezolve probleme din viața socială și economică.
	– Analizează oportunități de afaceri sau de dezvoltare antreprenorială, pornind de la cunoștințele dobândite în domeniul interfețelor de comunicare prin voce cu sisteme inteligente.
– Demonstrează abilități de management ale situațiilor din viața reală (de exemplu gestionarea corectă a timpului de învățare).	

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.)

Activitățile asistate prevăd (i) prezentarea succintă a conceptelor fundamentale cu care se va opera în ședința respectivă, (ii) explicarea detaliată a metodelor utilizate și a rezultatelor ce se doresc a fi obținute și (iii) realizarea, individual sau în echipe de doi studenți, a pașilor necesari proiectării, implementării sau evaluării sistemelor de recunoaștere a vorbitorului, după caz. În prima etapă, eminentă interactivă, sunt folosite metode clasice de predare (prelegerea și expunerea), având ca material suport prezentări electronice, materiale audio și/sau video. Sunt încurajate în permanență interacțiunile verbale cu studenții, întrebările, răspunsurile astfel încât demersul pedagogic să poată fi adaptat la posibilitățile de asimilare și învățare a studenților. În etapa a doua sunt explicați pașii practici pe care urmează să-i realizeze studenții și sunt descrise uneltele hardware și software și metodele științifice pe care aceștia trebuie să le utilizeze. În a treia etapă studenții implementează și evaluează independent problemele de studiu prin utilizarea calculatorului, a mediului software și a echipamentelor dedicate (atunci când este cazul).

Etapă a treia se extinde și în afara orelor propriu-zise, presupunând cel puțin 60% lucru individual suplimentar față de activitățile asistate. Aplicațiile realizate îi ajută pe studenți în dezvoltarea unor relații optime de comunicare într-un climat favorabil învățării prin descoperire. Materialele de proiect sunt disponibile studenților sub formă electronică pe platforma Moodle a facultății.

10. Conținuturi

PROIECT		
Nr. crt.	Conținutul	Nr. ore
1	Introducere în recunoașterea vorbitorului	4
2	Seturi de date pentru recunoașterea vorbitorului	4
3	Trăsături vocale obținute cu modele de învățare profundă preantrenate. Comparații, analize și evaluări subiective	4
4	Sisteme de recunoașterea vorbitorului bazate pe trăsături vocale obținute cu modele de învățare profundă preantrenate	4
5	Sumarizarea, prezentarea și interpretarea rezultatelor experimentale pentru diverse sarcini conexe recunoașterii vorbitorului	4
6	Evaluarea sistemelor de recunoașterea vorbitorului. Metrici de evaluare și analiza detaliată a erorilor	4
7	Proiectarea și antrenarea de la zero a unui sistem de recunoaștere a vorbitorului	4



	Total:	28
--	---------------	----

Bibliografie:

1. H. Cucu, *Proiect de cercetare-dezvoltare în tehnologia vorbirii: îndrumar de proiect*, disponibil în format electronic pe platforma Moodle a facultății de ETTI: <https://curs.upb.ro/>
2. D. Oneață, Al.L. Georgescu, H. Cucu, D. Burileanu, C. Burileanu, “Revisiting SincNet: An Evaluation of Feature and Network Hyperparameters for Speaker Recognition”, *Proceedings of the 28th European Signal Processing Conference (EUSIPCO)*, Amsterdam, the Netherlands, pp. 361-365, Dec. 2020, IEEE NY, ISBN: 978-9-0827-9705-3, DOI:10.23919/Eusipco47968.2020.9287794.
3. A. Buzo, H. Cucu, L. Petrică, D. Burileanu, *Metodă și sistem pentru diarizare în timp real a semnalelor audio, utilizate pentru recunoașterea automată a vorbirii și a vorbitorului*, brevet nr. 130883 B1 / 27.02.2019, înregistrat la OSIM.
4. V. Andrei, H. Cucu, C. Burileanu, “Overlapped Speech Detection and Competing Speaker Counting – Humans Versus Deep Learning”, *IEEE Journal of Selected Topics in Signal Processing*, vol 13, issue 4, pp. 850-862, Aug. 2019, IEEE NY, ISSN: 1941-0484, DOI:10.1109/JSTSP.2019.2910759, ISI IF 6.688.
5. Al.L. Georgescu, A. Caranica, H. Cucu, C. Burileanu, “RoDigits – a Romanian Connected-Digits Speech Corpus for Automatic Speech and Speaker Recognition”, *University “Politehnica” of Bucharest Scientific Bulletin, Series C*, Bucharest, vol. 80, issue 3, pp. 45-62, 2018, ISSN: 2286-3540.
6. Al.L. Georgescu, H. Cucu, “GMM-UBM Modeling for Speaker Recognition on a Romanian Large Speech Corpora”, *Proceedings of the 12th Romanian International Conference on Communications (COMM)*, Bucharest, Romania, 2018.
7. J.H.L. Hansen and T. Hasan, “Speaker Recognition by Machines and Humans: A Tutorial Review”, *IEEE Signal Processing Magazine*, 32(6):74–99, 2015.
8. M.W. Mak and J.T. Chien, “Tutorial on Machine Learning for Speaker Recognition”, *Proceedings of Interspeech*, San Francisco, USA, Sept. 2016.
9. J.S. Chung et al. “In Defence of Metric Learning for Speaker Recognition”, *Proceedings of Interspeech*, Shanghai, China, pp. 2977-2982, Oct. 2020, DOI: 10.21437/Interspeech.2020-1064.
10. N.R. Koluguri, T. Park, B. Ginsburg, “TitaNet: Neural Model for Speaker Representation with 1D Depth-wise Separable Convolutions and Global Context”, *Proceedings of the 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, pp. 8102-8106, Mai 2022.
11. J. Peng et al., “Improving Speaker Verification with Self-Pretrained Transformer Models”, *Proceedings of Interspeech*, Dublin, Ireland, pp. 5361-5365, Aug. 2023, DOI:10.21437/Interspeech.2023-453.

11. Evaluare

Tip activitate	11.1 Criterii de evaluare	11.2 Metode de evaluare	11.3 Pondere din nota finală
11.4 Curs	–	–	–



11.5 Seminar/laborator/proiect	Corectitudinea și complexitatea părții teoretice a proiectului. Cantitatea de experimente realizate și corectitudinea acestora. Modul de prezentare a rezultatelor.	Notarea acestei activități este în regim de verificare pe parcurs (cu notă de la 1 la 10). Se va realiza o evaluare a raportului de proiect și totodată o evaluare orală a cunoștințelor prin întrebări adresate studenților în urma prezentării raportului de proiect.	100%
11.6 Condiții de promovare			
<ul style="list-style-type: none">– Obținerea a 50% din punctajul total.– Crearea sistemului de recunoaștere a vorbitorului.– Analizarea calității sistemului de recunoaștere a vorbitorului.– Realizarea raportului de proiect.			

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)

Programa disciplinei oferă studenților suportul teoretic și practic necesar atât cunoașterii principiilor și tehnicilor de bază de recunoașterea vorbitorului, cât și proiectării și evaluării unor sisteme practice de recunoașterea vorbitorului, bazate pe prelucrarea semnalului vocal și învățare automată. 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 ce trebuie să realizeze, din motive de securitate, și identificarea vorbitorului.

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 vorbitorului), în realizarea unor interfețe interactive sau a unor sisteme evaluate de dialog cu autentificare vocală.

Data completării

Titular de curs

Titular(i) de aplicații

–

Conf. dr. ing. Horia CUCU



Universitatea Națională de Știință și Tehnologie Politehnica București
Facultatea de Electronică, Telecomunicații și
Tehnologia Informației



Data avizării în departament

Director de departament

31.10.2024

Prof. Dr. Claudiu DAN

Data aprobării în Consiliul Facultății

Decan

01.11.2024

Prof. Dr. Mihnea Udrea