

შესავალი ხელოვნურ ინტელექტში

კურსის სილაბუსი

სასწავლო კურსის სტატუსი	საბაზისო სასერტიფიკატო კურსი BSCTBC108
შეხვედრებისა და საათების რაოდენობა	ხანგრძლივობა: 3 თვე შეხვედრების რაოდენობა: 20 ლექცია საათების რაოდენობა: 40 საათი
ლექტორი	ქეთევან მაჭავარიანი Email: ketevan.machavariani@geolab.edu.ge ნათია ფორჩხიძე Email: natia.porchkhidze@geolab.edu.ge
სასწავლო კურსის ფორმატი	თეორიული და პრაქტიკული ლექციები მოიცავს: ინტერაქციულ ლექციას, ინდივიდუალურ და ჯგუფურ სამუშაოებს, დისკუსიას, პრაქტიკულ სამუშაოს. კურსი დაყოფილია 4 ბლოკად. თითოეული ბლოკი შედგება 3 ლექციისა და ერთი შუალედური პროექტისგან. კურსის ბოლოს იქნება დამოუკიდებელი სამუშაო პერიოდი ფინალური პროექტისთვის მოსამზადებლად.

#	სასწავლო კურსის შინაარსი	ძირითადი მიზნობები
ბლოკი 1	შესავალი ხელოვნურ ინტელექტში (AI)	
ლექცია 1 (2 საათი)	რა არის ხელოვნური ინტელექტი - AI-ის ისტორია და ევოლუცია - ხელოვნური ინტელექტის ტიპები - ხელოვნური ინტელექტის გამოყენება ყოველდღიურ ცხოვრებაში - AI ეთიკა და უსაფრთხოება დავალება 1: დაწერეთ ერთგვერდიანი ესე, ერთ-ერთი თქვენთვის საინტერესო AI აპლიკაციის შესახებ. ისაუბრეთ რა პრობლემის გადაჭრისათვის გამოიყენება და აღწერეთ როგორ მუშაობს იგი.	- ხელოვნური ინტელექტის, ტიპები და გამოყენების სფეროები. - მანქანური სწავლების საფუძვლები და მოდელების შექმნა. - მონაცემების კრიტიკული როლი და მოდელის შეფასება.

ლექცია 2
(2 საათი)

მანქანური სწავლის საფუძვლები

- განსხვავება AI-ს, ML-სა და DL-ს შორის
- ML-ის სხვადასხვა ტიპები
- როგორ იწვრთნება მანქანური მოდელები.

დავალება #2: შექმენით მარტივი გადაწყვეტილებების ხე (decision tree) ყოველდღიური მაგალითების გამოყენებით (მაგ., "უნდა წავიღო ქოლგა დღეს?"). გარდა ამისა, მოიძიეთ ინტერნეტში და ახსენით მარტივი სიტყვებით, როგორ მუშაობს სარეკომენდაციო სისტემა (მაგ., YouTube, Facebook-ის სარეკომენდაციო სისტემა).

ლექცია 3
(2 საათი)

მონაცემების მნიშვნელობა ხელოვნურ ინტელექტში

- მონაცემთა სახეობები, მათი შეგროვება და დამუშავება
- კავშირი მონაცემთა ხარისხს, მონაცემთა რაოდენობასა და მოდელის სიზუსტეს შორის
- სხვადასხვა მოდელების შეფასების მეტრიკები

შუალედური დავალება #1:

მოამზადეთ პრეზენტაცია (5-7 გვერდიანი) შემდეგ პროექტზე:

1. შეარჩიეთ რეალურ სამყაროში არსებული ერთ-ერთი პრობლემა, რომლის მოგვარებაც შესაძლებელია ხელოვნური ინტელექტის გამოყენებით.
 2. გამოიკვლიეთ პრობლემა და იფიქრეთ როგორ შეიძლება გამოყენებულ იქნას AI ან მანქანათმცოდნეობა ამ პრობლემის მოსაგვარებლად. რა ტიპის ხელოვნური ინტელექტის ან მანქანათმცოდნეობის მოდელი შეიძლება იყოს გამოყენებული (მაგ. კლასიფიკაცია, რეგრესია, ა.შ.).
 3. მოიძიეთ წყაროები, სადაც შეგიძლიათ მიიღოთ მონაცემები თქვენი პროექტისთვის (მაგ., საჯაროდ ხელმისაწვდომი მონაცემთა ნაკრები, გამოკითხვები ან მარტივი ექსპერიმენტები).
 4. ახსენით, როგორ მოხდება მონაცემების დამუშავება
 5. ისაუბრეთ რომელი მეტრიკის გამოყენებით შეაფასებდით მოდელის სიზუსტეს.
- პრეზენტაცია სასურველია შეიცავდეს ვიზუალურ ელემენტებს (მაგ., დიაგრამები ან სქემები), სადაც ეს შესაძლებელია, რათა უკეთ ახსნათ თქვენი პროექტი. დაგეხმაროთ თქვენი პროექტის ახსნაში.

ლექცია 4 (2 საათი)	შუალედური პროექტი 1-ის წარდგენ	
ბლოკი 2	პროგრამირება პითონით (PYTHON) ხალოვნური ინტელექტისთვის	
ლექცია 5 (2 საათი)	დავალება 3 შეასრულეთ შემდეგი დავალებები Google Colab-ში: • დაწერეთ პითონის სკრიპტი რომელიც: დაბეჭდავს "გამარჯობა [თქვენი სახელი და გვარი]!" ეკრანზე. • დაწერეთ პითონის ფუნქცია, რომელიც იღებს ორ რიცხვს, აჯამებს მათ და აბრუნებს შედეგს. • დაწერეთ პითონის ფუნქცია, რომელიც ითვლის მართკუთხედის ფართობს მისი სიგრძისა და სიგანის გათვალისწინებით. • დაწერეთ ფუნქცია რომელიც გაარკვევს შემოსული სიტყვა პალინდრომია თუ არა (ერთნაირად იკითხება წინიდან და უკნიდან, მაგალითად: hannah, noon, mad-am, etc.) • ბონუსი: დაწერეთ პითონის ფუნქცია, რომელიც რომაულ ციფრებს გადაიყვანს არაბულ ციფრებში (მაგ. გადასცემთ string ფორმატში XII დაგიბრუნებთ int ფორმატში 12-ს)	• პროგრამირება პითონის გამოყენებით და კოდის წერის საფუძვლები. • მონაცემებთან მუშაობა და მათი მანიპულირება • პითონის ბიბლიოთეკებისა და API-ების გამოყენება მონაცემთა ანალიზისთვის
ლექცია 6 (2 საათი)	მონაცემებთან მუშაობა პითონით • სიები, ლექსიკონები და Numpy მასივები • ფაილების წაკითხვა • Pandas-ის გამოყენება მონაცემთა მანიპულირებისთვის. დავალება #4: შეასრულეთ შემდეგი დავალებები Google Colab-ში: • დაწერეთ პითონის ფუნქცია, რომელსაც გადავცემთ რიცხვების სიას და დაგვბრუნებს არსებული რიცხვების საშუალოს და მედიანას. გააკეთეთ pandas და numpy ბიბლიოთეკების მეშვეობით. კომენტარად დაუწერეთ, რა უპირატესობა აქვს არსებული ბიბლიოთეკების გამოყენებას. • დაწერეთ პითონის სკრიპტი, რომელიც წაიკითხავს CSV ფაილს Pandas-ს მეშვეობით. დაპრინტეთ dataframe-ის პირველი ხუთი ხაზი. შემდეგ, გაიგეთ dataframe-ის სვეტების ტიპები, დაპრინტეთ მარტივი ანალიტიკა ამ სვეტებისთვის (mean, std, min, max... describe მეთოდით), გამოიტანეთ თითოეული ცვლადისთვის ობზერვაციების რა პროცენტია ნაკლები და ბოლოს წაშალეთ მწკრივები ნაკლები მნიშვნელობებით.	

ლექცია 7 (2 სკატი)

პითონის ბიბლიოთეკებთან და API-ებთან მუშაობა

- პითონის ბიბლიოთეკების გამოყენება
- ურთიერთქმედება API-ებთან
- მონაცემთა ვიზუალიზაცია Python-ით
- კლასიფიკაციის ამოცანის მაგალითის განხილვა

შუალედური დავალება 2:

1. დარეგისტრირდით [kaggle.com](https://www.kaggle.com)-ზე
2. ჩამოტვირთეთ ტიტანიკის დაცა <https://www.kaggle.com/competitions/titanic>
3. წაიკითხეთ მონაცემები პითონით
4. გაწმინდეთ და დაამუშავეთ მონაცემები Pandas-ს გამოყენებით
5. დაბეჭდეთ ანალიტიკა და 2-3 რელევანტური გრაფიკი (მაგ. სამიზნე ცვლადის (გადარჩენის) განაწილება ასაკისა თუ ბილეთის კლასის მიხედვით)
6. sklearn -ის მეშვეობით დაატრენინგეთ რომელიმე (ან რამდენიმე) tree based მოდელი (random forest, xgboost, catboost)
7. შეაფასეთ მოდელი გავლილი მეტრიკებით (confusion matrix, accuracy, precision, recall, roc-auc curve, f1)
8. ბონუსი: გამოიტანეთ მოდელებისათვის shap ამხსნელი ცვლადები

ლექცია 8 (2 სკატი)

შუალედური პროექტი 2-ის წარდგენა

ბლოკი 3	ბუნებრივი ენის დამუშავება (NLP)	
ლექცია 9 (2 საათი)	NLP საფუძვლები - როგორ აანალიზებს მანქანური მოდელი ტექსტურ მონაცემებს - შესავალი Hugging Face და Transformers - წინასწარ დატრენინგებული ენის მოდელების გამოყენება დავალეზა 5 დანერეთ პითონის ფუნქცია, რომელიც: - დაგიყოფს წინადადებას სიტყვებად. - დაგიბრუნებს უნიკალური სიტყვების სიას. - მოახდენს უნიკალური სიტყვების სიხშირის ვიზუალიზაციას - ბონუსი: დანერეთ პითონის ფუნქცია, რომელსაც გადასცემთ სიტყვას და დაგიბრუნებთ მის დამარცვლულ ვერსიას (მაგ. Input: გამარჯობა, output: გა,მარ,ჯო,ბა)	- ბუნებრივი ენის დამუშავების (NLP) საფუძვლები და ტექსტური მონაცემების ანალიზი - NLP მოდელების პრაქტიკული გამოყენება და მათი შეფასება

ლექცია 10 (2 საათი)	ტექსტურ მონაცემებთან მუშაობა - ტექსტის მონაცემების წაკითხვა და წინასწარი დამუშავება - ტექსტის ვექტორული წარმოდგენა - Input და Output ფორმატების გაგება NLP მოდელებისთვის დავალეზა 6 შექმენით წინადადების ვექტორული წარმოდგენა. - დანერეთ პითონის ფუნქცია, რომელიც იღებს წინადადებას და აბრუნებს სიტყვების სიას (ტოკენები). ამოიღეთ სასვენი ნიშნები. - დანერეთ ფუნქცია, რომელიც აქცევს მოცემულ ტექსტს ლექსიკონად, სადაც keys არის სიტყვები და values არის მათი სიხშირე. - დანერეთ ფუნქცია, რომელიც დააბრუნებს ლექსიკონს, სადაც მოცემულ წინადადებაში თითოეული უნიკალური სიტყვა აისახება ვექტორულად. - განახორციელეთ მარტივი TF-IDF და word2vec წარმოდგენა და შეამოწმეთ იგი მარტივ წინადადებაზე.	
---------------------------------------	--	--

ლექცია 11
(2 საათი)

NLP მოდელების გამოყენება

- წინასწარ დატრენინგებული მოდელების გამოყენება ტექსტის კლასიფიკაციისთვის
- მოდელის სიზუსტის შეფასება

შეაღწეოთ დავალება 3

შექმენით სენტიმენტების ანალიზის ინტერაქციული ინსტრუმენტი წინასწარ გაწვრთნილი NLP მოდელის გამოყენებით და წარმოადგინეთ მარტივი აპლიკაცია Streamlit ან Gradio-ს გამოყენებით.

- გამოიყენეთ წინასწარ გაწვრთნილი NLP მოდელი (მაგ., Hugging Face Transformers) განწყობის ანალიზის შესასრულებლად.
- მიეცით საშუალება მომხმარებლებს შეიყვანონ რეალურ სამყაროში არსებული ტექსტური მონაცემები და გამოიყენონ მოდელი სენტიმენტების კლასიფიკაციისთვის (დადებითი, უარყოფითი ან ნეიტრალური).
- შექმენით მოსახერხებელი ინტერფეისი Streamlit-ით ან Gradio-ით, სადაც მომხმარებლებს შეუძლიათ შეიყვანონ ტექსტი და ნახონ განწყობის შედეგები დინამიკურად.
- დატესტეთ როგორც ინგლისურ ტექსტზე დატრენინგებული რომელიმე მოდელი, ასევე multilingual მოდელი (ქართულისათვის).
- კომენტარად მიუწერეთ თქვენი აღმოჩენა: რა განსხვავებაა ამ მოდელების პერფორმანსს შორის? რომელს აქვს თქვენი აზრით უკეთესი პერფორმანსი და რატომ?

ლექცია 12
(2 საათი)

შეაღწეოთ პროექტი 3-ის წარდგენა

ბლოკი 4	მანქანური ხედვა (CV)	
ლექცია 13 (2 საათი)	შესავალი კომპიუტერულ ხედვასა და გამოსახულების დამუშავებაში • კომპიუტერული ხედვის საფუძვლები • პიქსელებიდან მახასიათებლებამდე: როგორ აანალიზებს კომპიუტერები გამოსახულებებს • სურათების დამუშავება: ფილტრაცია, ტრანსფორმაციები და კიდეების ამოცნობა დავალება 7 დანერეთ პითონის სკრიპტი, რომელიც: • წაიკითხავს სურათს. • გარდაქმნის სურათს ნაცრისფერ ფერში. • დაბურავს სურათს და ამოიცნობს კიდეებს. • წარმოადგენს ვიზუალიზაციას ორიგინალური სურათისა და დაგენერირებული შედეგების გვერდიგვერდ.	• კომპიუტერული ხედვის (CV) საფუძვლები და გამოსახულების დამუშავება • CV მოდელების პრაქტიკული გამოყენება და მათი შეფასება
ლექცია 14 (2 საათი)	გამოსახულებით მონაცემებთან მუშაობა • სურათის მონაცემების ჩატვირთვა და ტრანსფორმაცია • სურათის აუგმენტაციის ტექნიკები • Huggingface CV მოდელების გამოყენება დავალება 8 დანერეთ სკრიპტი, რომელიც: • წაიკითხავს სურათს • შეასრულებს სურათის რამდენიმე ტრანსფორმაციას. ასევე, მოიძიეთ ინფორმაცია და გამოიკვლიეთ როგორ მუშაობს თვითმართვადი მანქანები.	

ლექცია 15 (2 საათი)	შუალედური დავალება 4 შექმენით გამოსახულების კლასიფიკაციის ინსტრუმენტი, რომელიც განსაზღვრავს ძაღლების ჭიშებს წინასწარ გაწვრთნილი მოდელის გამოყენებით და განათავსეთ იგი Streamlit ან Gradio-ით. - გამოიყენეთ წინასწარ გაწვრთნილი ღრმა სწავლების მოდელი ძაღლების ჭიშების კლასიფიკაციისთვის. - ნება მიეცით მომხმარებლებს ატვირთონ ძაღლების სურათები და გამოიყენონ მოდელი მათი ჭიშის პროგნოზირებისთვის. - განავითარეთ მოსახერხებელი ინტერფეისი Streamlit ან Gradio-ით, სადაც მომხმარებლებს შეეძლება ატვირთონ სურათები და ნახონ კლასიფიკაციის შედეგები. - დაამატეთ ნდობის ქულები, მრავალი სურათის ატვირთვა ან მოდელის პროგნოზირების ალბათობების ვიზუალიზაცია (ბონუსი)	
ლექცია 16 (2 საათი)	შუალედური პროექტი 4-ის წარდგენა	
ფინალური ბლოკი	ფინალურ პროექტზე მუშაობა და ლექტორთან უკუკავშირი	
ლექცია 17 (2 საათი)	პროექტების განხილვა/პრეზენტაცია და ლექტორთან უკუკავშირი	- AI ტექნიკების გამოყენება. - მონაცემებზე დაფუძნებული AI მოდელის შექმნა რეალური პრობლემების გადასაჭრელად. - პროექტის პრეზენტაცია
ლექცია 18 (2 საათი)	პროექტების განხილვა/პრეზენტაცია და ლექტორთან უკუკავშირი	
ლექცია 19 (2 საათი)	პროექტების განხილვა/პრეზენტაცია და ლექტორთან უკუკავშირი	
ლექცია 20 (2 საათი)	პროექტების განხილვა/პრეზენტაცია და ლექტორთან უკუკავშირი	