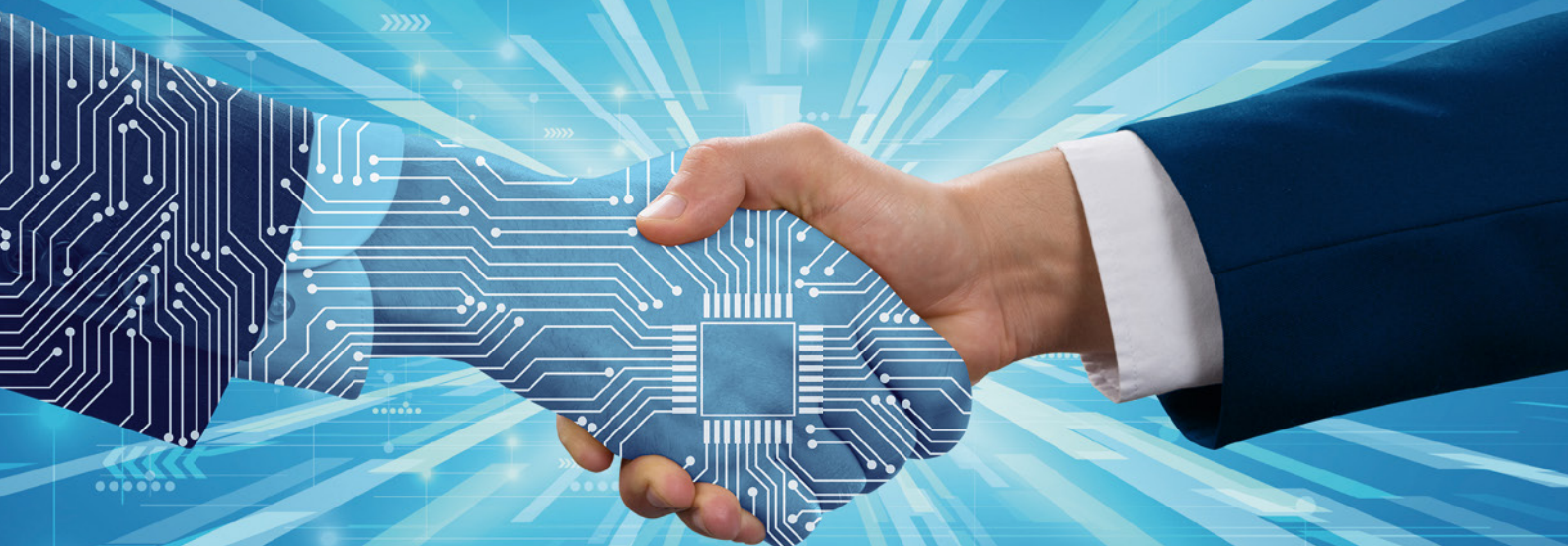




Data Science E^xpert

300H UČENJA SA TRENERIMA
SRPSKI / ENGLESKI
UČIONICA / ON-LINE
POJEDINAC / IN-HOUSE
RADNO VREME / NERADNO VREME



MENADŽMENT
CENTAR
BEOGRAD



CILJ

Naučite data science od osnova Python-a do naprednog mašinskog učenja i steknite sve veštine da biste pospešili svoju karijeru.

Od biblioteke Pandas do dubokog učenja, posle završenog kursa ćete znati kako da istražujete, čistite i transformišete podatke u praktična zapažanja i kako da implementirate modele mašinskog učenja od početka do kraja u proizvodno okruženje, radeći u timovima i sa našim profesionalnim trenerima.

CILJNA GRUPA

Svi zaposleni koji se bave **analizom podataka**: analitičari, planeri, izveštavanje, controlling, FP&A, CRM, business intelligence, data science, data analitika, itd.

Svi ostali koji žele da započnu karijeru kao:

- **junior data scientist** za neku od najboljih srpskih IT kompanija (HTEC, Nordeus, Vega IT, enjoy.ing, sa.global itd.)
- **data analitičar, data inženjer, ili data menadžer**
- preduzetnici pokretanjem usluge obrade podataka ili proizvodom, osnivanjem **startup-a**, prikupljanjem novca kako biste ubrzali rast i upravljali svojim data science timom

NAČIN RADA

- Kompletan program traje **300h** u periodu od 3 do 6 meseci.
- Obuka može biti **online** (Zoom, Teams) ili u učionici.
- Obuka može biti na **srpskom** jeziku ili engleskom.
- Naši **treneri** su sve vreme prisutni i rade sa vama.
- Imate pristup našoj **e-learning platformi** na kojoj su sačuvani svi **materijali** koji se koriste tokom rada, kao i **snimci** online sastanaka.
- Vreme je **fleksibilno** i postoji mogućnost rada u toku ili van radnog vremena, radnim i neradnim danima.
- Obuka može da se radi za pojedinca ili kao **in-house** za kompaniju

SADRŽAJ

1. ANALIZA PODATAKA U PYTHON-U

- Osnove programiranja u Python-u
- Python za Data Science
- Baze podataka
- Relacione baze podataka i SQL
- Vizualizacija podataka

2. MATEMATIKA I STATISTIKA ZA DATA SCIENCE

- Statistika
- Verovatnoća
- Linearna algebra i analiza

3. MAŠINSKO UČENJE I DUBOKO UČENJE

- Prethodna obrada podataka
- Nadzirano učenje
- Nenadzirano učenje
- Duboko učenje
- Ocena i optimizacija modela
- Generalizacija

4. DATA SCIENCE PROJEKTI

- Organizacija projekta
- Google Cloud Platforma (GCP)
- Grupni projekat

1. ANALIZA PODATAKA U PYTHON-U

Naučićete osnove programiranja u Python-u, kako da radite s Jupyter Notebook-om i Jupyter Lab-om i upoznaćete se s moćnim Python bibliotekama koje se koriste u data science, kao što su Pandas i NumPy, za istraživanje velikih setova podataka i obavljanje statističkih analiza.

Osnove programiranja u Python-u

Upoznaćete se s vrstama podataka i promenljivama, uslovima, petljama, funkcijama, klasama i strukturama podataka (npr. lista i rečnik).

Python za Data Science

Naučićete programiranje u Python-u, kako da radite s Jupyter Notebook-om i da koristite moćne Python biblioteke, kao što su Pandas i NumPy, da istražujete i analizirate velike skupove podataka.

Baze podataka

Prikupljate podatke iz raznih izvora, uključujući CSV fajlove, SQL upite za relacione baze podataka, Google Big Query, API-je i Web scraping.

Relacione baze podataka i SQL

Naučićete kako da formulišete dobar SQL upit. Upoznaćemo se sa arhitekturom šeme relacione baze, a potom zalazimo duboko u naprednu manipulaciju funkcijom SELECT kako bismo izvukli korisne informacije iz baze podataka.

Vizuelizacija podataka

Upoznaćemo se sa najboljim načinima da prenesete ključne poruke vizualizacijom podataka. Naučićete kako da grafički prikazete podatke pomoću Python biblioteka kao što su Matplotlib i Seaborn i da transformišete podatke u praktična zapažanja, koristeći ispravne vizuelizacije.



2. MATEMATIKA I STATISTIKA ZA DATA SCIENCE

Steći ćete razumevanje matematike koja leži u osnovi svih biblioteka i modela za analizu podataka i mašinsko učenje. Upoznaćete se s osnovnim konceptima statistike i verovatnoće, kao i s linearnom algebrom, u srži numeričkih operacija u bibliotekama kao što su Pandas i NumPy.

Statistika

Upoznaćete se sa pojmom slučajnih promenljivih, raspodelama, srednjim vrednostima, varijansama itd. Naučićete kako da testirate hipoteze (A/B testiranje). Upoznaćete se sa Python bibliotekom statsmodels za statističku obradu podataka.

Verovatnoća

Upoznaćete se sa osnovnim konceptima verovatnoće koji podupiru statistiku i omogućavaju njeno razumevanje, kao što su centralna granična teorema, Bajesova teorema, slučajni događaji, uslovna verovatnoća, itd.

Linearna algebra i analiza

Ovladaćete radom sa matricama (vektorima), koji su u osnovi operacija u Python bibliotekama kao što su Pandas i NumPy. Produbićete znanje vezano za analizu matematičkih funkcija, izvode i slično, kako biste razumeli princip rada algoritama mašinskog učenja.



3. MAŠINSKO UČENJE I DUBOKO UČENJE

Samostalno ćete kreirati modele mašinskog i dubokog učenja koji će rešavati probleme poput predviđanja cena, klasifikacija, klasterovanja, detekcije izuzetaka, razumevanja fotografija, obrade prirodnog jezika i slično.

Prethodna obrada podataka

Naučite kako da istražujete, čistite i pripremate set podataka pomoću tehnika kao što su standardizacija, odabir karakteristika, inženjering karakteristika i slično.

Nadzirano učenje

Naučite kako da rešavate zadatke predviđanja i klasifikacije pomoću scikit-learn biblioteke koristeći algoritme kao što su linearna i logistička regresija, KNN, SVM, metode ansambla (bagging, bosting, stacking), slučajne šume itd.

Nenadzirano učenje

Implementiraćete metode kao što su PCA za smanjenje dimenzionalnosti ili klasterovanje za otkrivanje grupa u podacima.

Duboko učenje

Razumećete arhitekturu neuronskih mreža (neurona, slojeva, stekova) i njihovih parametara (aktivacija, gubitaka, optimizatora). Postaćete samostalni u izradi sopstvenih mreža, učeći tehnike i trikove koji omogućavaju duboko učenje. Upoznaćete se sa bibliotekom Tensorflow Keras.

Ocena i optimizacija modela

Poboljšavaćete performanse modela pomoću metoda kao što su unakrsna validacija ili hiperparametarska optimizacija. Pretvorićete tok rada mašinskog učenja u ponovljiv niz koraka (pipeline).

Generalizacija

Naučićete kako da pomoću metoda regularizacije sprečite preterano prilagođavanje podacima i kako da izaberete funkciju gubitka da biste poboljšali preciznost modela.



4. DATA SCIENCE PROJEKTI

Zaokružićete svoje data science znanje pakovanjem svog koda u funkcionalnu aplikaciju, dostunu korisnicima. Upoznaćete se sa tehnikama organizacije projekta i koda, kao i alatima za pravljenje projekta i kolaboraciju sa ostalim kolegama.

Organizacija projekta

Naučićete kako da pravilno postavite projekat mašinskog učenja da biste brzo i sa sigurnošću vršili iteracije. Naučićete kako da model mašinskog učenja konvertujete u model s robustnim i skalabilnim pipeline-om pomoću biblioteke sklearn-pipeline koristeći kodere i transformatore. Koristićemo biblioteku MLflow za pravljenje toka rada, pravićemo sopstveni API, Docker kontejner i primenićemo API na Cloud Run-u. Izradićemo front-end vebsajta na kom će se pokretati aplikacija.

Google Cloud Platforma (GCP)

Implementiraćete kod i model u proizvodno okruženje. Pomoću Google Cloud AI platforme, moći ćete da trenirate svoj model u velikom obimu (skaliranje), da ga upakujete i učinite dostupnim svima. Koristićete skladišni prostor u cloud-u (Cloud Storage) i virtuelne mašina za pokretanje koda. Koristićete Docker okruženje da implementirate RESTful Flask API, koji se može ubaciti u bilo koji front-end interfejs.

Grupni projekat

Kroz saradnju u malim timovima, iskoristićete sve veštine koje ste naučili kako biste rešili realan data science problem. Upoznaćete se sa radom sa GitHub-om i deljenjem koda. Koristićete kombinacije sopstvenih setova podataka (ako ih imate od svoje kompanije) i repozitorijuma otvorenih podataka (vladine inicijative, Kaggle itd.). Primenićete u praksi sve alate, tehnike i metodologije obrađenih na ovom treningu. Smišljavaćete, prezentovaćete, dizajniraćete, programiraćete i implementiraćete potpuno funkcionalni data science projekat koji će se može koristiti u realnim situacijama i kao aplikacija spremna za portfolio radi prikazivanja vašeg novog seta data science veština.



BROJ 1 ZA DATA SCIENCE



Ustanička 189, 11050 Beograd
+381 11 3047 126, +381 63 8500 991
office@mcb.rs, www.mcb.rs



#mcbbeograd