

# Mašinsko učenje

## Drugi domaći zadatak

24.05.2020

Tema drugog domaćeg zadatka su akcije na berzi, konkretno [S&P 500](#) indeks.

**Prvi deo zadatka (10 poena)** zahteva kreiranje rekurentne neuralne mreže (RNN) sa LSTM ćelijama korišćenjem Tensorflow biblioteke za predviđanje cene akcije S&P 500 indeksa.

Na sledećem linku možete naći najnovije podatke sa [Yahoo Finance](#). Možete uzeti proizvoljni vremenski interval od nekoliko meseci/godina u kome ćete posmatrati cene akcija. Ukoliko koristite Yahoo Finance dovoljno je da posmatrate **daily close price** u vašem modelu.

Od vas se očekuje da pokrijete sve vezano za ovaj problem: pronalaženje, čišćenje i priprema podataka, kreiranje modela, treniranje i testiranje modela.

Za prikazivanje rezultata možete koristiti biblioteku po izboru, bitno je vizuelno prikazati odnos stvarne i predviđene cene.

**Drugi deo zadatka (15 poena)** takođe podrazumeva kreiranje rekurentne neuralne mreže koja će ovog puta raditi sa akcijama pojedinačnih kompanija iz S&P 500 indeksa, a ne samog indeksa.

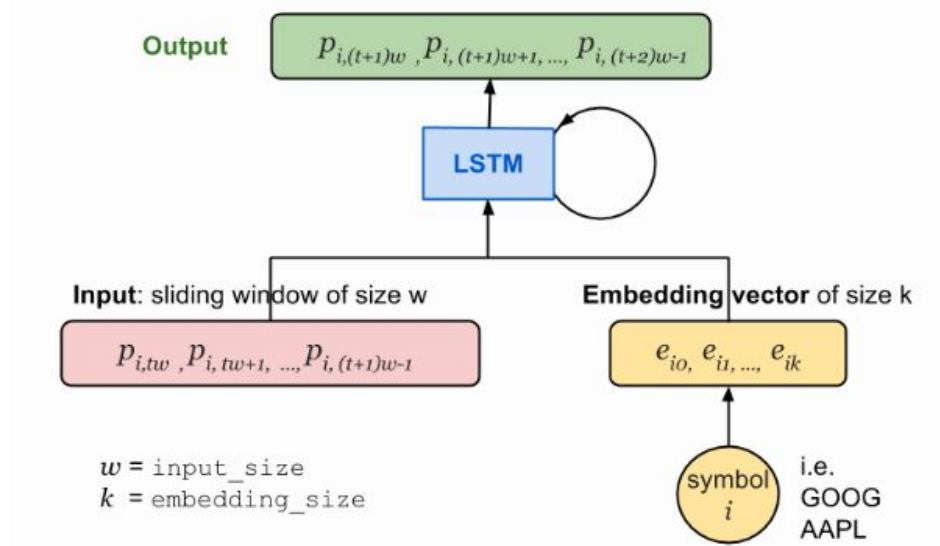
Podatke možete nabaviti na [nekoliko načina](#), bitno je da izaberete 50-100 kompanija i radite sa istorijskim podacima njihovih akcija u istom vremenskom intervalu.

Ovde je takođe dovoljno da koristite **daily close price** kao feature.

Kako želimo da model nauči ponašanje pojedinačnih akcija, koje će imati drugačije pattern-e moramo mu reći sa kojom akcijom radi. Ovde treba da postignete korišćenjem [embedding-a](#). Ulaz u RNN u trenutku  $t$  sadrži vektor koji je rezultat konkratenacije vektora cene  $i$ -te akcije veličine  $w$  kao i embedding simbola akcije (AAPL, GOOG). Slika ilustruje ovaj deo sistema.

Rezultat ovog dela domaćeg treba takođe da bude vizuelizacija, samo ovog puta grafikon treba da prikaže stvarnu i predviđenu cenu za svaku kompaniju koju ste posmatrali.

Poželjno je da napravite funkciju koja kao argument uzima simbol akcije (GOOG, AAPL) i iscrtava traženi grafik (nemojte crtati 100 grafikona).



Pri izradi domaćeg zadatka moguća je saradnja studenata u grupama od **najviše troje**. Pritom, svaki student će biti nezavisno ocenjivan na osnovu odbrane pri kojoj se proverava razumevanje predatog rešenja i relevantnog gradiva. Iako je podela posla dozvoljena, ukoliko jedan član tima ne razume neke delove predatog rešenja, za te delove će mu biti dodeljeno 0 poena. Ukoliko postoji deo predatog rešenja koji niko iz tima ne razume, svim članovima tima će biti dodeljeno 0 poena na celom domaćem zadatku.

Domaći zadatak se izrađuje i predaje na sledeći način:

1. Možete koristiti Colab Notebook, Jupyter Notebook ili raditi lokalno u editoru po želji. Struktura projekta je na vama.
2. Ceo projekat zapakovati u fahivu **ml\_d2\_x\_y\_z** gde su x, y i z prezimena članova tima. Nemojte slati datasetove koje koristite.
3. Slanje fajla **ml\_d2\_prezime1\_prezime2\_prezime3** na mejl adresu [iciganovic@raf.rs](mailto:iciganovic@raf.rs) ili [mmilunovic@raf.rs](mailto:mmilunovic@raf.rs) u zavisnosti od toga kod koga slušate vežbe. Subject mejla mora biti oblika "[ML D2] prezime1 prezime2 prezime3". U tekstu mejla obavezno navesti članove tima sa imenom, prezimenom i brojem indeksa.

**Rok za slanje rešenja je 04.06.2020 23:59**

U slučaju nepoštovanja bilo kog od navedenih pravila (naslov mejla, ime arhive, rok za slanje) rad neće biti pregledan i svim članovima tima će biti dodeljeno 0 poena.

**Odbrana domaćeg zadatka će biti održana na poslednjim vežbama 05.06.2020.**

Raspored odbrana će biti objavljen naknadno. Ukoliko predate rešenje ranije i više vam odgovara odbrana pre početka roka, možemo organizovati dodatni termin.