

The background of the slide is a histological micrograph showing various cellular structures stained with hematoxylin and eosin (H&E). The image displays clusters of cells with prominent nuclei (stained blue/purple) and surrounding tissue with pinkish-red staining, likely representing connective tissue or extracellular matrix. The overall texture is granular and complex, typical of biological tissue sections.

TIM PANDE

Bjanka Bašić, Ivan Knezić i Jelena Lončar

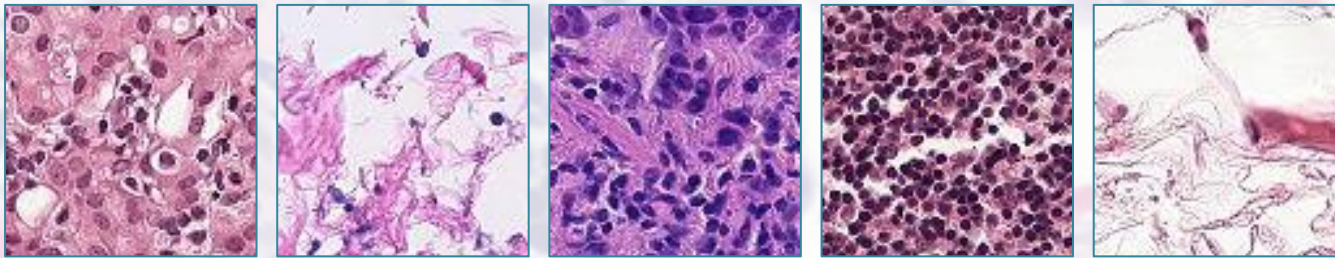
Detekcija raka na temelju histoloških preparata

Opis problema

- Problem: detekcija prisutnosti metastatskog raka na temelju fotografija histoloških preparata tkiva
- Dijagnostička procedura dugotrajna je i naporna za patologe
 - potrebno ispitati veliko područje tkiva
 - malene se metastaze lako mogu previdjeti
- Svodi se na problem **binarne klasifikacija** – oznaka klase (y) može poprimiti samo dvije vrijednosti
 - $y=1$ – pozitivan primjer
 - $y=0$ – negativan primjer
- Naš zadatak: izgraditi binarni klasifikator za prepoznavanje tumornog tkiva

Opis podataka

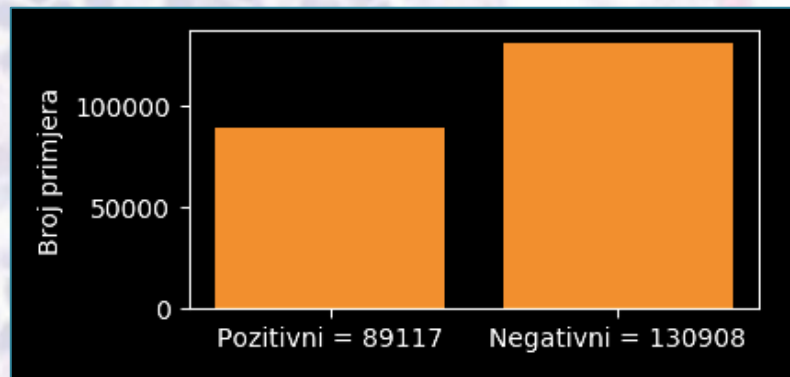
- Podaci: fotografije(96x96px) koje predstavljaju mikroskopske snimke tkiva limfnih čvorova, na koje je primjenjeno HE bojenje



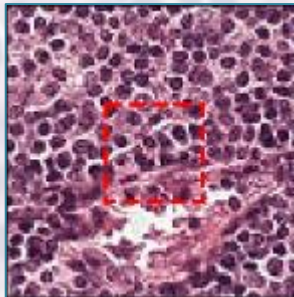
- Primjer se smatra pozitivnim primjerom ako centralno područje fotografije dimenzija 32x32px sadrži barem jedan piksel tumorskog tkiva – okolno područje ne utječe
- Skup podataka kombinacija je podatak prikupljenih:
 - Radboud University Medical Center (Nijmegen, Nizozemska)
 - University Medical Center Utrecht (Utrecht, Nizozemska)
- Koristimo verziju sa Kaggle-a - ne sadrži duplikate

Eksploratorna analiza - zaključci

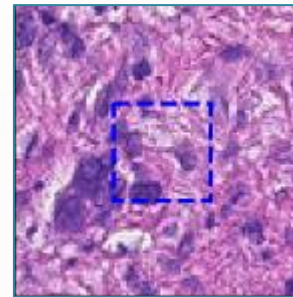
- U datasetu se nalazi: 220025 primjera za treniranje
57458 primjera za testiranje
- Omjer negativnih i pozitivnih je **59/41**



- Negativni primjer

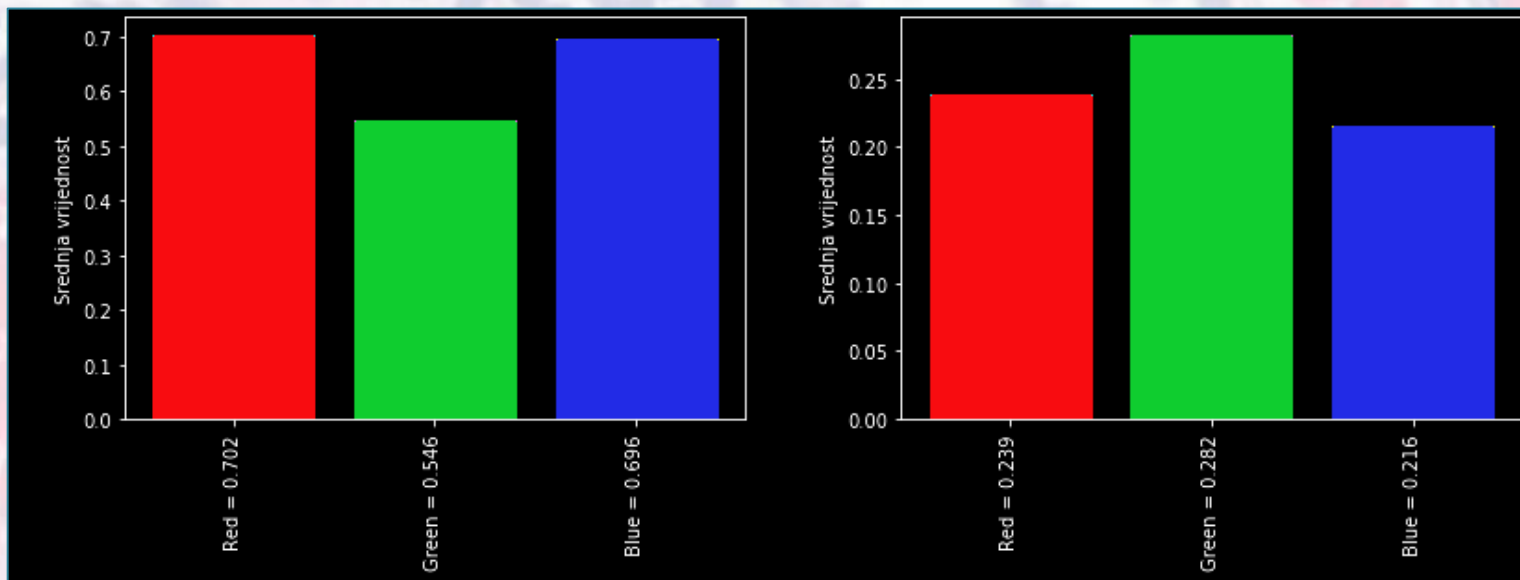


- Pozitivni primjer



Eksploratorna analiza - zaključci

- Statistička analiza: srednje vrijednosti pojedinih kanala boje



- 7 primjera presvjetlih/pretamnih slika za zaključiti nešto iz njih – zanemarivo na sveukupan broj train primjera

Cilj

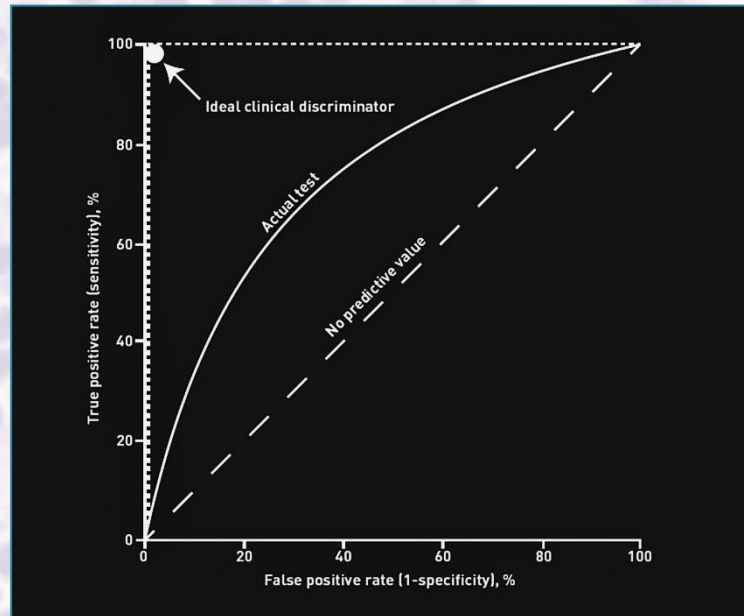
- Cilj: primjenom adekvatnih algoritama strojnog učenja postići da model bude što bolji prediktor za vrijednost ciljne varijable odgovarajućeg elementa prostora primjera
- Za dani primjer, odnosno fotografiju, odrediti što točnije je li oznaka klase:
 - $y=1$ – na slici se nalazi barem jedan pixel tumorskog tkiva
 - $y=0$ – na slici se ne nalazi tumorsko tkivo

Metodologija

- Problem ćemo rješavati primjenom metoda i algoritama nadziranog učenja – **konvolucijske neuronske mreže (duboko učenje)**
- Nastojati ćemo dobiti što uspješniji model istraživanjem raznolikih mogućnosti za arhitekture KNM te hiperparametre
- Planiramo:
 - Koristiti Python i razne biblioteke za njega:
 - Numpy, Randas, Scikit-learn, Keras, Fast.ai
 - Kao razvojno okruženje koristiti Jupyter
 - Na fotografije (primjere) primijeniti augmentacijske transformacije

Evaluacija modela

- Pozitivni primjeri su nam važniji od negativnih pa ćemo evaluirati sa – **površinom ispod ROC krivulje**



- Uz to još ćemo odrediti matricu konfuzije te konkretne vrijednosti osjetljivosti, preciznosti te F1-mjere

The background of the slide is a histological micrograph of tissue, likely stained with hematoxylin and eosin (H&E). It shows various cellular structures, including clusters of cells with dark purple nuclei and pinkish-red cytoplasm and extracellular matrix. The overall texture is granular and complex, typical of biological tissue sections.

Hvala na pažnji!