NEODATA **HACKATHON**

**ZADATAK**  
Vaš zadatak je koristiti satelitske slike preuzete s Google Mapsa kako biste naučili model za klasifikaciju piksela u dvije kategorije: **cesta (road)** i **pozadina (background)**. Cilj je razviti model koji će precizno identificirati i razdvojiti ceste od pozadine analizirajući slike i učeći iz danih podataka. Svaki piksel slike mora biti označen etiketom:

* **1** za cestu
* **0** za pozadinu

Cilj je razviti algoritam za segmentaciju cesta na temelju danih podataka. Drugim riječima, vaš model mora što preciznije razlikovati cestu od ostalih objekata na satelitskoj slici.

**Podaci**

1. **training.zip**

* Trening set koji sadrži slike i njihove odgovarajuće "ground truth" maske. maski koje definiraju što je cesta (1) i što je pozadina (0). Ovaj skup podataka omogućuje modelu da nauči obrasce potrebne za segmentaciju cesta.
* Satelitske slike prikazuju stvarne scenarije (gradove, sela, ceste I prirodu) :

An aerial view of a neighborhood

Description automatically generated

* Ground truth maske (binarne slike) koje služe kao referentne vrijednosti za učenje. Svaki piksel maske sadrži:
  + **(bijela boja) : Cesta**

1. **(crna boja) : Pozadina**

Maska (training/groundtruth):   
A black and white grid

Description automatically generated

**Dimenzije** slika i njihovih maski su identične. To osigurava da se piskeli točno podudaraju s pikselima odgovarajuće ulazne slike.

Trenirate Vaš model na training setu, pritom pazeći da se broj slike iz images foldera poklapa s brojem slike iz groundtruth foldera.

1. **test\_set\_images.zip – testni set**

* Testni set slika na kojima se evaluira vaš model.
* Testni set je skup slika koji se koristi za **evaluaciju vašeg modela**. Ove slike sadrže neoznačene podatke koje vaš model treba obraditi kako bi predvidio maske za segmentaciju cesta.
* Nema priložene "ground truth" maske jer se te maske koriste interno od strane organizatora za evaluaciju.
* Vaš model treba generirati izlazne maske za svaku sliku iz testnog seta.
* Izlazne maske pretvaraju se u CSV datoteku prema pravilima formata za predaju.
* NAPOMENA: Obavezno provjerite formate i dimenzije maski za predaju, jer se predaja pogrešnih formata neće uzeti u obzir i bodovat će se s 0.

1. **Sample\_submission.csv**

* Datoteka sample\_submission.csv služi kao uzorak formata kakav trebate predati pri evaluaciji.
* Prvi stupac: **ID slike** te koordinate gornjeg lijevog kuta "patcha" (dijela slike) dimenzija **16x16 piksela**.
* Format ID-ja je: ID\_x\_y gdje:
  + - ID označava identifikator slike
    - X i Y označavaju koordinate gornjeg lijevog kuta patcha unutar slike.
    - Primjer: 1\_16\_32 znači da je ID slike jednak 1 te da patch počinje na koordinatama (16,32)
* Drugi stupac sadrži oznaku, tj labelu za taj patch:
* 1: patch sadrži cestu
* 0: patch sadrži pozadinu
* NAPOMENA: **Obavezan format predaje:** Predana datoteka mora **slijediti isti format** kao sample\_submission.csv.

Svaka pogreška u formatu (npr. zaboravljene zareze, pogrešni ID-ovi, izostavljeni redci) može rezultirati odbijanjem predaje.

* Nakon što vaš model napravi predikcije za svaki piksel slike, trebate ih grupirati u **16x16 piksela**.
* Generirajte oznake za svaki patch (npr prosječna vrijednost piksela unutar patcha)

1. **Mask\_to\_submission.py**

* Ova Python skripta služi za konverziju binarnih maski (slike s označenim pikselima za ceste i pozadinu) u CSV datoteku koja je spremna za predaju u sklopu hackathona. To se radi pretvaranjem svake maske u niz predikcija za svaki patch dimenzija 16x16 piksela.
* Skripta koristi fiksnu veličinu patcha od **16x16 piksela**

1. **Submission\_to\_mask**

* Omogućava vizualizaciju rezultata predikcija modela tako što pretvara numeričke predikcije iz CSV datoteke u binarne maske.

**EVALUACIJA**

Vaš algoritam se ocjenjuje prema F1 ocjeni (kombinira dvije metrike: preciznost i osjetljivost).

Napravite account na sljedećoj stranici putem vašeg FER ili PMF maila:

<https://www.aicrowd.com/challenges/epfl-ml-road-segmentation>

Odaberite opciju Participate.A screenshot of a computer

Description automatically generated

**RESURSI**

Svi resursi potrebni za sudjelovanje u hackathonu dostupni su kao open-source i slobodno ih možete koristiti.

Savjet: koristiti google colab s opcijom T4 GPU.

**PRAVILA**

1. Troje sudionika po jednom timu. Imenujte tim.
2. Svaki član tima može napraviti račun te ima 5 pokušaja za predaju. NEMOJTE odabrati opciju „create Team“ jer će tada vaš tim imati ukupno 5 mogućih predaja umjesto 15. Prijavite se zasebno te tada sveukupno imate 15 pokušaja za predaju i ocjenjivanje. Pažljivo predajte pazeći pritom na format csv datoteke.
3. Najbolja predaja (najbolje ocjenjena) je ujedno i ona koju će organizatori uzeti u obzir prilikom ocjenjivanja i rangiranja. Treba ju pokazati organizatorima prilikom završavanja zadatka.
4. Na kraju hackathona sudionici su obavezni poslati svoj **kod** i **CSV datoteku** koja će se koristiti za predstavljanje njihovog rješenja tijekom završne prezentacije. Mail na koji šaljete: [neodata@eestec.hr](mailto:neodata@eestec.hr)
5. Sudionici moraju pripremiti i predstaviti **pitch** kako bi pokazali što su postigli tijekom hackathona. Njihova prezentacija bit će dio ukupne ocjene i evaluirat će se na temelju jasnoće i kvalitete rješenja koje su razvili. Prezentirat će se na zatvaranju radionice.
6. Svako kršenje pravila biti će kažnjeno.

Sretno!  
  
NeoData tim