

# Korisnička dokumentacija

## Instalacija

Projektno rješenje pisano je u programskom jeziku Python 3 koristeći čitav niz pomoćnih biblioteka (Tensorflow, scikit-learn, scikit-image...).

U direktoriju s izvornim kodom nalazi se datoteka `requirements.txt` koja sadrži popis potrebnih biblioteka.

Preporuka je da koristite virtual environment za njihovu instalaciju.

Instalacija virtualenv:

```
pip3 install virtualenv
```

Stvaranje novog virtualnog okruženja

```
virtualenv -p /usr/bin/python3 mozgalo
```

Aktivacija stvorenog virtualnog okruženja

```
source ./mozgalo/bin/activate
```

Instalacija svih potrebnih paketa:

```
pip3 install -r <korijenski_direktorij_projekta>/requirements.txt
```

---

## Skupovi podataka

Korišteno je nekoliko javno dostupnih skupova podataka za evaluaciju rješenja:

- [cifar10](#)
- [stl10](#)
- [Cats\\_vs\\_dogs](#)

Kako bi pojednostavnili skidanje i pohranu svih skupova podataka potrebnih za reproduciranje rezultata pripremili smo jednu arhivu koja sadrži sve potrebno na sljedećoj [poveznici](#). Dobivenu arhivu raspakirati u korijenski direktorij projekta.

```
wget wget -O dataset.tar.gz 'https://www.dropbox.com/s/7ui2a58md0zvhep/dataset.tar.gz?dl=0'
tar xvf dataset.tar.gz
```

---

## Pokretanje

U direktoriju **notebooks** nalazi se jupyter bilježnica s cjelovitim rješenjem.

---

```
cd notebooks
```

```
jupyter notebook final_report.ipynb
```

Odlučili smo se za korištenje python bilježnice radi jednostavnosti korištenja, pogodnosti za demonstraciju rezultata, laganog uređivanja i eksperimentiranja nad rezultatima.

U samoj bilježnici slijedno je prikazan postupak:

- učitavanje podataka
- izlučivanje značajki
- odabir optimalne dimenzije značajki i redukcija dimenzionalnosti
- odabir optimalnog broja grupa i samo grupiranje
- prikaz rezultata
- provjera rezultata nad označenim skupovima podataka

---

## Struktura direktorija

- **dataset** - direktorij sa skupovima podataka
- **notebooks** - korištene Jupyter bilježnice
- **models** - prethodno trenirani TensorFlow modeli korišteni za ekstrakciju značajki
- **report\_imgs** - slike korištene u dokumentaciji
- **clusters** - dobiveni rezultati grupiranja
- **src** - direktorij s izvornim kodom
- **src/modules/external** - git submodule iz paketa **slim** i

**keras** za korištenje gotovih prethodno treniranih mreža

---

Priloženo programsko rješenje je isprobano pod Linux okruženjem (Ubuntu 16.04 i Arch Linux).

Ako naiđete na ikakav problem prilikom pokretanja, molimo Vas da se javite nekom od članova tima kako bi otklonili moguće poteškoće.