PREDVIĐANJE BRZINE VJETRA   
Projektni prijedlog

Ivan Avirović, Antonio Đurić, Mario Marjanović

UVODNI OPIS PROBLEMA

Problem koji ćemo pokušati riješiti je predviđanje brzine vjetra pomoću dobivenih atributa pomoću meteoroloških senzora.  
Dani atributi su dnevne prosječne oborine, maksimalna i minimalna dnevna temperatura, minimalna temperatura trave te nekoliko indikatora.

Skup podataka koje koristimo smo preuzeli sa stranice Kagglea [[1].](https://www.kaggle.com/datasets/fedesoriano/wind-speed-prediction-dataset)

Skup podataka se sastoji od 6574 dnevnih mjerenja, od 1.1.1961. do 31.12.1978., koristeći 5 meteoroloških senzora ugrađenih u meteorološkoj stanici. Uređaj koji je prikupljao podatke se nalazio na značajno praznom području.

Precizno predviđanje brzine vjetra može biti od iznimne ekonomske važnosti (npr. agrikultura, vjetroelektrane) te može pomoći u sprječavanju prirodnih katastrofa.

CILJ I HIPOTEZE ISTRAŽIVANJA PROBLEMA

Za početak ćemo zanemariti činjenicu da se radi o vremenskom nizu, te ćemo pokušat predvidjeti brzinu vjetra koristeći sve značajke osim datuma.

Pretpostavljamo da su određeni uvjeti, npr. niska minimalna dnevna temperatura zraka i minimalna temperatura trave, ukazuju na veću brzinu vjetra.

Naknadno ćemo uzeti datum u obzir, te ćemo ponovno pokušati predvidjeti brzinu vjetra koristeći sve značajke. Jer na brzinu vjetra dosta često utječe brzina vjetra nekih prijašnjih dana i na primjer godišnje doba, očekujemo da će ove predikcije biti točnije.

PREGLED DOSADAŠNJIH ISTRAŽIVANJA

Postoji popriličan broj istraživanja koji se bavi predviđanje brzine vjetra.

[[2]](https://www.codespeedy.com/wind-direction-and-speed-prediction-using-machine-learning-in-python/) koristeći drugi skup podataka sličnim atributima objašnjava osnovne mogućnosti neuronskih mreža strojnih učenja.

[[3]](https://arxiv.org/pdf/2007.12567.pdf) također koristeći drugi skup podataka i višedimenzionalne konvolucijske neuronske mreže pokušava se predvidjeti brzina vjetra

[[4]](https://www.mdpi.com/1996-1073/15/4/1510/pdf) pokazuje kako se može iskoristiti napredak u dubokom učenju u primjeni na obnovljivim izvorima energije između ostalog i vjetra.

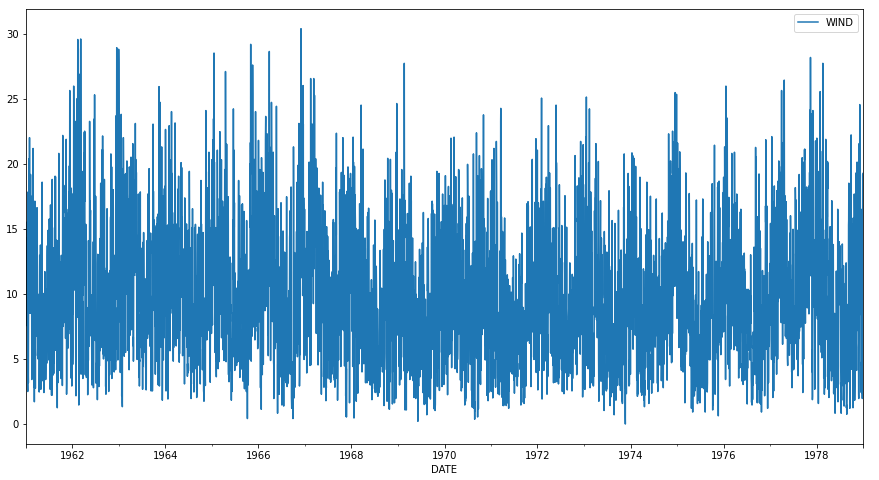
Na Kaggleu na istom skupu podataka trenutno postoji manja analiza podataka osobe koja je iste te podatke objavila prije nekoliko dana.

MATERIJALI, METODOLOGIJA I PLAN ISTRAŽIVANJA

Sav kod napisat ćemo u Pythonu u Jupyter bilježnicama koje ćemo stavljati na stranicu <https://github.com/iavirovic/strojno-ucenje>.

Jer nam neki featuri nedostaju u podacima, isprobat ćemo razne metode data imputationa poput: mean imputation, regression imputationa itd. Također ćemo napraviti feature scaling kako sama razlika u veličinama značajki ne bi utjecala na rezultate modela.

Iz sljedećeg grafa vidimo da je brzina vjetra sezonalna(na početku godine načelno bude više):



Metode, algoritmi i tehnike koje ćemo koristiti:

* Linearna regresija
* SVR
* ARIMA
* XGBoost
* ...

Uspješnost rezultata našeg modela koji ne koristi datume ocijeniti ćemo train & test metodom ostavljajući trećinu početnog skupa podataka za testiranje te k-strukom unakrsnom validacijom. Kod modela koji koristi datume moramo biti pažljiviji, jer ne želimo model trenirati na podacima iz budućnosti kako oni ne bi utjecali na predikcije. Isprobat ćemo neke od metoda iz [[7].](https://medium.com/@soumyachess1496/cross-validation-in-time-series-566ae4981ce4)

OČEKIVANI REZULTATI PREDLOŽENOG PROJEKTA

Očekujemo da će naš model sa što većom točnošću predviđati brzinu vjetra temeljeno na zadanim atributima. Također očekujemo da će modeli koji uzimaju datum u obzir biti točniji od onih koji zanemaruju datum.

LITERATURA

[1] <https://www.kaggle.com/datasets/fedesoriano/wind-speed-prediction-dataset>

[2] <https://www.codespeedy.com/wind-direction-and-speed-prediction-using-machine-learning-in-python/>

[3] <https://arxiv.org/pdf/2007.12567.pdf>

[4] <https://www.mdpi.com/1996-1073/15/4/1510/pdf>

[5] <https://github.com/pmf-strojnoucenje/Vjezbe>

[6] <https://web.math.pmf.unizg.hr/nastava/su/>

[7] <https://medium.com/@soumyachess1496/cross-validation-in-time-series-566ae4981ce4>