

Kapacita turbín veterných elektrární

Kapacita turbíny je maximálne množstvo elektrickej energie, ktoré môže turbína vyprodukovať za ideálnych podmienok (napríklad pri optimálnej rýchlosti vetra). Vyjadruje sa v kilowattoch (kW) alebo megawattoch (MW). Výkon turbín v kW vo veterných elektrárňach závisí od mnohých faktorov (napríklad výška turbíny v metroch, priemer rotora v metroch a ďalšie). V tejto úlohe budeme túto tému skúmať na väčšej dátovej sade ([uloha3_data.xlsx](#)) veterných elektrární, tentokrát v Spojených štátoch amerických¹.

V nami pripravenom súbore ([uloha3_data.xlsx](#)) sme vybrali niektoré atribúty dát týkajúce sa kapacity turbín veterných elektrární. Vaše úlohy sú nasledujúce:

a) Pomocou vhodne zvoleného postupu vypočítajte alebo zistite, ktorý parameter z nasledujúcich najviac súvisí s kapacitou turbíny (t_{cap}): priemer rotora v metroch (t_{rd}), výška turbíny v metroch (t_{hh}) alebo súradnica zemepisnej dĺžky (y_{lat})? Vzťah parametrov s výkonom taktiež vizualizujte.

¹ Nasledujúce webstránky vám pomôžu sa v dátach zorientovať – s datasetom sa môžete zoznámiť v interaktívnej vizualizácii, metadáta zas obsahujú technické informácie o atribútoch dát.

Vizualizácia: <https://eerscmap.usgs.gov/uswtdb/viewer/#5.3/37.501/-117.751>

Metadáta: https://eerscmap.usgs.gov/uswtdb/assets/data/uswtdb_v7_1_20240814.xml

b) V tejto časti bude vašou úlohou pracovať s vopred pripraveným zdrojovým kódom (notebookom) v Pythone (**uloha3_pomocka.ipynb**). Nemusíte sa ničoho báť, kód nebude potrebné písať, postačí iba načítať a spustiť pripravený súbor.

Otvorte si prostredie Google Colab (<https://colab.research.google.com/>) a cez tlačidlo **Nahrat'** pridajte súbor **uloha3_pomocka.ipynb**. Súbor s dátami **uloha3_data.xlsx** si môžete nahráť na Google Disk. Otvorený zdrojový kód spustíte pomocou tlačidla **Runtime** a **Spustiť všetko**. Google Colab si ešte môže vyžiadať potvrdenie o tom, že súhlasíte s prístupom na Google Disk.

V notebooku **uloha3_pomocka.ipynb** je pripravených niekoľko metód strojového učenia, ktoré predikujú kapacitu jednotlivých turbín alebo ktoré vedia rozhodnúť, v akom štáte sa daná turbína nachádza.

Spustite si vopred pripravený zdrojový kód v nástroji Google Colab pomocou predchádzajúceho návodu a zodpovedajte na nasledujúce otázky.

- Aké metódy strojového učenia sú použité na riešenie úloh v notebooku?
- Akú úspešnosť dosiahli jednotlivé modely strojového učenia v jednotlivých úlohách?
- Čo je to tréningová a testovacia množina dát?
- Aký je rozdiel medzi klasifikačnými a regresnými metódami?
- Čo je to hyperparameter? Čo sú to one-hot vektory a ako ich využívame v našom notebooku?

Ak sa Vám na niektoré z otázok nepodarí zodpovedať, nebojte sa, v ďalšom kole, si viaceré z týchto otázok vysvetlíme na jednoduchých príkladoch.

Poznámky pre riešenie úloh prvého kola:

Pri riešení úloh **môžete používať internet :)**, pri niektorých úlohách to bez neho ani nepôjde. Voľba softvéru je na vás, či už využijete tabuľkový softvér ako MS Excel a jeho alternatívy alebo sa vydáte cestou programovania, či využitia iných nástrojov na prácu s dátami, mapami atď. V jednej časti budete musieť aj programovať, a to v ľubovoľnom programovacom jazyku.

S prípadnými otázkami sa na nás môžete kedykoľvek obrátiť. Riešenia úlohy (dokumentácia + prípadný zdrojový kód) môžete odovzdať v termíne do **31.01.2025** cez formulár, ktorý je zverejnený na stránke <https://vucap-challenge.science.upjs.sk/>

Riešenia jednotlivých podúloh vhodne okomentujte, ak je to vhodné pridajte aj obrázky. Je možné odovzdať aj čiastočné riešenia jednotlivých úloh. Pri veľmi zaujímavom či prepracovanom riešení (pod)úlohy vám môžu byť udelené aj bonusové body.
