**Žilinská univerzita v Žiline**

**Fakulta riadenia a informatiky**

**Data imputation using PCA**

*Semestrálna práca 2*

**Úvod do strojového učenia**

**David Kučera**

Akademický rok 2024/2025

Obsah

[Popis práce 3](#_Toc186279074)

[Stručný popis programu 3](#_Toc186279075)

[Prezentácia a zhodnotenie výsledkov výpočtových experimentov 3](#_Toc186279076)

# Popis práce

Druhá semestrálna práca sa zameriava na implementáciu a výpočtové experimenty s vybranou metódou strojového učenia – Data imputation using PCA.

# Stručný popis programu

Program bol implementovaný v programovacom jazyku **Python** vo vývojovom prostredí JetBrains PyCharm. Obsahuje jeden **main.py** súbor, v ktorom sa nachádzajú všetky zdrojové kódy programu.

Obsahuje nasledovné funkcionality:

* načítanie vstupných dát zo súboru
* deskriptívna štatistika dát
* doplnenie chýbajúcich dát strednou hodnotou
* výpočet MAE, RMSE
* doplnenie chýbajúcich dát pomocou PCA z knižnice sklearn
* doplnenie chýbajúcich dát pomocou vlastne implementovanej PCA
* vlastná implementácia metódy PCA
* štandardizácia dát
* ...

# Prezentácia a zhodnotenie výsledkov výpočtových experimentov