



Trojanova 13, 120 00 Praha 2

### Posudek vedoucího bakalářské práce

**Autor práce:** Jan Trödler  
**Název práce:** Metody strojového učení ve fyzice pevných látek  
**Vedoucí práce:** doc. Jan Vybíral

#### Téma předkládané práce:

Předkládaná bakalářská práce studenta Jana Trödlera se zabývá možnostmi využití metod strojového učení na datech generovaných v současné době v oboru fyziky pevných látek. Metodami simulace vlastností nejrůznějších materiálů se zabývá stále větší počet pracovišť na celém světě a jejich zpracování metodami strojového učení umožňuje tato data ještě lépe využít ke konstrukci nových materiálů. Data použitá v předkládané práci pocházejí ze spolupráce s kolegy z Fritz-Haberova institutu v Berlíně a týkají se klíčových parametrů materiálů potencionálně vhodných pro konstrukci solárních panelů. Před několika lety byla použita v jedné z otevřených soutěží serveru [kaggle.com](https://www.kaggle.com), kde bylo úkolem účastníků předpovídat formační energii a tzv. *bandgap energy* krystalů hliníku, gallia a india.

#### Splnění požadavků zadání:

Student se seznámil se strukturou poskytnutých dat a jejich fyzikálním významem. Dále nastudoval základní metody strojového učení a možnosti jejich aplikace na dostupná data. Ve své práci rekapituluje matematické základy metody *Kernel Ridge Regression* včetně odvození metody nejmenších čtverců a rozboru její kombinace s použitím jader. Podařilo se mu jak reprodukovat stávající algoritmy zpracování těchto dat, tak i navrhnout vlastní modifikace těchto algoritmů. Tím student vesměs splnil zadání práce. Tyto modifikace nebyly nicméně dostatečné na to, aby byla výrazně překonána současná úroveň predikce důležitých materiálových vlastností dosažená již v soutěži organizované na [kaggle.com](https://www.kaggle.com).

**Obsahová úroveň práce:**

Vlastní práce je členěna do tří kapitol. První kapitola podává souhrn použitých metod, včetně potřebného matematického základu a jejich odvození. Ve druhé kapitole jsou popsána zpracovávaná data, jejich základní charakteristiky a dělení na trénovací, testovací a validační data. Konečně třetí kapitola podává přehled o použití metody kernel ridge regression na různé reprezentace dostupných dat.

**Vlastní práce studenta:**

Při zpracování této práce musel student zvládnout řadu dílčích úkolů. Nastudoval teoretické základy použitých metod strojového učení a jeho výklad svědčí o dobré orientaci v této problematice. Dále se seznámil s implementací metody kernel ridge regression v programu Matlab a strukturou předložených dat. A konečně navrhl několik způsobů aplikace metod strojového učení na různé reprezentace vstupních dat. Použité postupy a dosažené výsledky nicméně nepřesahují významným způsobem publikaci [12], jak plyne i ze srovnání v tabulce 3.7.

**Závěr:** Bakalářskou práci doporučuji k obhajobě a navrhuji klasifikovat známkou

**velmi dobře (B).**

V Praze, 18. 8. 2020

doc. Jan Vybíral