

Strojové učení pro středoškoláky

V rámci projektu plánujeme zorganizovat pěti nebo šestitýdenní večerní kurz strojového učení pro zájemce z pražských a dojezdově dostupných středočeských gymnázií. Studenty se pokusíme srozumitelně a zajímavě (ale přesto věcně správně) seznámit s tím, co je to strojové učení, jaké je jeho využití v praxi a proč by se něčím takovým mohli chtít zabývat.

Chceme tak zvýšit povědomí o strojovém učení a o tom, že je možné se jím na Matfyzu zabývat. Domníváme se, že drtivá většina studentů středních škol si pod studiem informatiky představí zejména programování (pokud si tedy vůbec něco představí). To je škoda, protože informatika toho zahrnuje mnohem více, a právě strojové učení považujeme za velice zajímavou a zároveň užitečnou oblast. Částečně neprávem je strojové učení považované za něco velice složitěho a nepřístupného, a běžně se s ním nesetká dokonce ani bakalářský student Matfyzu; význam tohoto oboru přitom ukazuje fakt, že figuruje v názvech dvou magisterských studijních zaměření.

Přitom například *naivní Bayesův klasifikátor* je myšlenkově triviální, pro pochopení vyžaduje pouze znalost podmíněné pravděpodobnosti, která je běžnou součástí středoškolských osnov, a lze jej snadno implementovat i v tabulkovém procesoru, nebo dokonce manuálně s pomocí papíru a tužky. Strojové učení umožňuje řešení mnoha praktických úloh, které jsou jinými způsoby jen obtížně řešitelné. Studium strojového učení zároveň umožňuje porozumět technologiím, které se stávají běžnou součástí našeho života, jako je například rozpoznávání řeči, spam-filtry, řazení zobrazených příspěvků na Facebooku, ...

Plánovaný kurz zorganizujeme v období *říjen až prosinec 2015*. Studenti budou na kurz docházet jednou týdně do budovy MFF UK na Malostranském náměstí. Většinu času našich dvouhodinových „lekci“ budou tvořit přednášky. Studenti budou také dostávat domácí úkoly na programování. Ty s námi budou konzultovat především prostřednictvím e-mailu. Závažnější problémy budeme řešit přímo na hodinách. Předběžná navrhovaná struktura kurzu je následující:

- ukázat a vysvětlit, co je strojové učení
- *perceptron, množina rysů*, geometrická intuice (*separační nadrovina*)
- úvod do statistiky a pravděpodobnosti
- *logistická regrese*
- *naivní Bayesovský klasifikátor*
- *přeučení, normalizace hodnot rysů*

V současné době jednáme s *OMK* o zveřejnění plánovaných učebních textů na serveru *matfyz.cz*. Přestože tyto texty budou svým rozsahem značně přesahovat obsah kurzu, poslouží jako základ pro plánované přednášky. Zpětnou vazbu, kterou v rámci kurzu získáme, bychom rádi využili k vylepšení těchto textů a k získání lepší představy, jaká programovací cvičení jsou pro takovéto studenty odpovídající. K propagaci akce bychom rádi využili kontakty organizátorů jiných akcí propagujících Matfyz, jako je ŠMF, korespondenční semináře nebo olympiády.

Při loňském projektu *Propaguj svou školu na své škole*, kdy jsme pořádali přednášky přímo na středních školách, jsme zjistili, že drtivá většina středoškoláků nejeví o studium informatiky zájem. Proto jsme usoudili, že lepších výsledků dosáhneme, když potenciální zájemce o studium pozveme přímo na půdu fakulty.

Na projekt žádáme částku *9 000 Kč*. Z toho *1 000 Kč* na propagační a výukové materiály ke kurzu a *8 000 Kč* na stipendia pro účastníky projektu (Mgr. Jindřich Libovický, Mgr. Rudolf Rosa, Bc. Tomáš Musil), které rozdělíme podle míry zapojení do projektu.