

NÁVRH A REALIZACE TRADING STRATEGIÍ NA BÁZI STROJOVÉHO UČENÍ S POMOCÍ MATLABU

RNDr. Miroslav Pavelka, PhD

m.pavelka@sh.cvut.cz

Ing. Jan Hovad

jan@hovad.cz

OBSAH

- × Obchodování a strojové učení
- × Specifika prediktivního modelování u burzovních časových řad
- × Redefinice targetu pro účely obchodování
- × Přehled použitých modelů, výhody a nevýhody
- × Návrh systému v Matlabu
- × Několik poznámek o Matlabu ve vztahu implementaci trading strategií
- × Závěr
- × Literatura a odkazy

OBCHODOVÁNÍ A STROJOVÉ UČENÍ

✘ Obchodování

- + Dynamická úloha $\xrightarrow[\text{(vstupní podmínky)}]{\text{Vstup}} \xrightarrow[\text{(výstupní podmínky nebo sl.)}]{\text{Výstup}}$
- + TA indikátory

✘ Strojové učení

- + Nejobvyklejší případ:

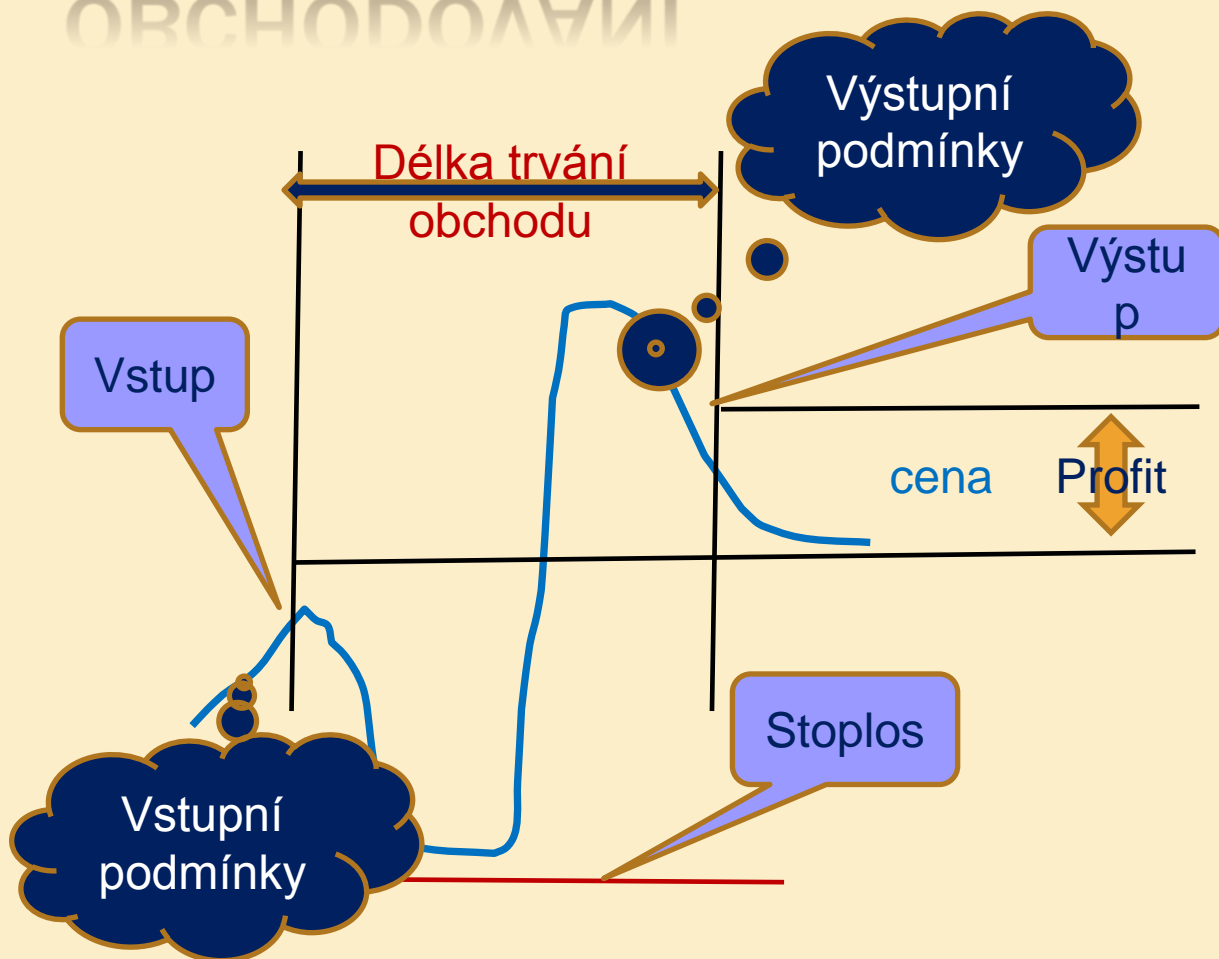
$$X \rightarrow y, X = (x^1 \dots, x^D), y = y^1$$

- + Učení s učitelem, učení bez učitele (shlukování)
- + Aplikace v mnoha oborech, data mining

✘ Data mining a obchodování

- + Desítky tisíc prací
- + Většinou: fixní časový interval (do následujícího dne příp. periody, diskrétní target, cenový vývoj do následujícího dne)
- + Většinou EOD data

OBCHODOVÁNÍ



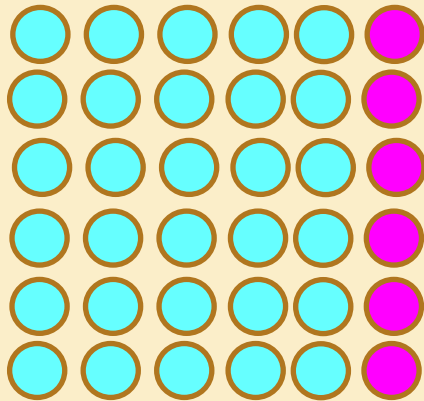
- × Jak formalizovat?
- × Optimal stopping theory
- × Úloha sekretářky

- × Dva vzájemně spojené definující body
- × Vstup (ve spojitém čase) //Vstupní podmínky
- × (Různě) dlouhá doba trvání obchodu
- × Výstup (ve spojitém čase) //Výstupní podmínky
- × Velká flexibilita, přímé uplatnění TA
- × Nelze převést na úlohu prediktivního modelování

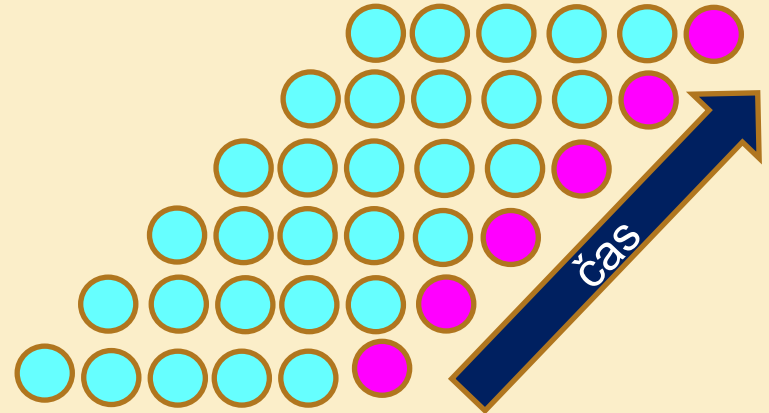
SPECIFIKA PREDIKTIVNÍHO MODELOVÁNÍ U BURZOVNÍCH ČASOVÝCH ŘAD

- ✘ Správná definice targetu je základ (pokud se nejedá o akademickou studii bez vazby na praxi)
- ✘ Mezi proměnnými je časová vazba, nutná „mrtvá oblast“ mezi trénovacími a validačními daty
- ✘ Data nejsou stacionární, nutné inkrementální učení nebo časté přeučování
- ✘ Koncept drift je problém, Ensemble modelů s různě dlouhou trénovací množinou
- ✘ Odměna za uhodnutí targetu a pokuta za chybu se významně liší i v rámci jedné skupiny targetů
- ✘ Snadná tvorba korelovaných multidimenzionálních dat (zvláště pokud použijeme Hankelovy matice), obtížná selekce proměnných
- ✘ Nutno několik úrovní testování

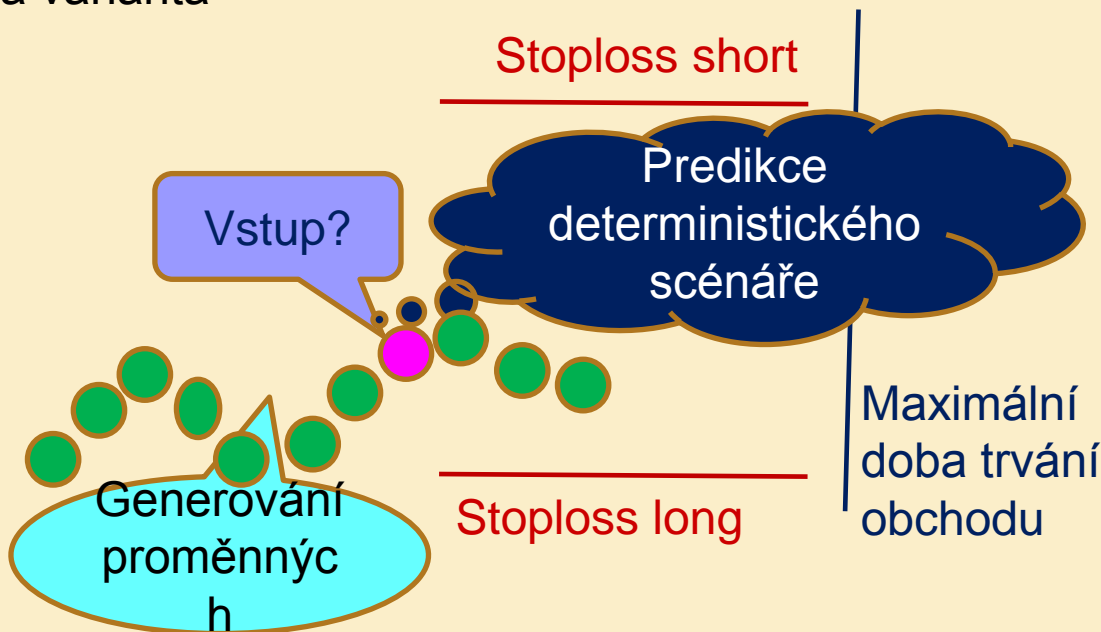
REDEFINICE TARGETU PRO ÚČELY OBCHODOVÁNÍ



Statická varianta



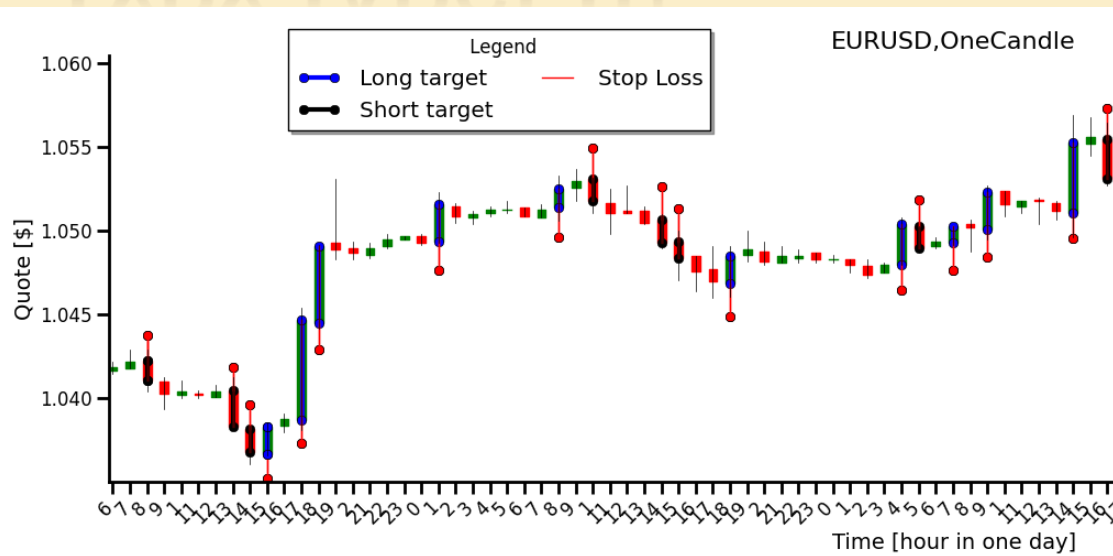
dynamická varianta



TYPY TARGETU A TVORBA PROMĚNNÝCH

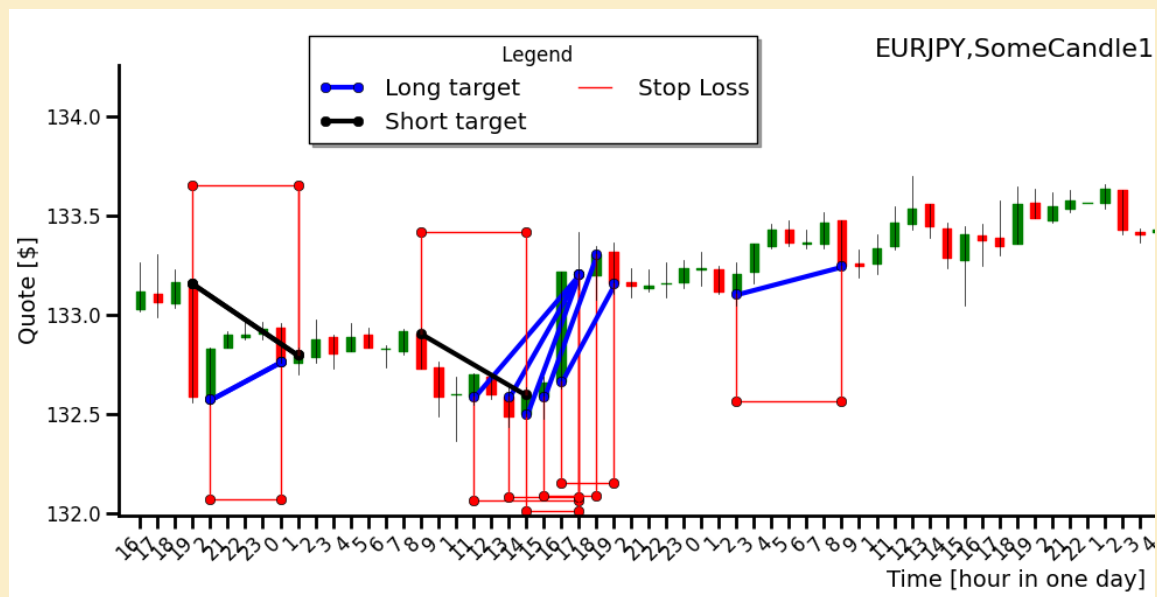
- ✘ Zatím 4 typy targetu:
 - + Onecandle
 - + Somecandle
 - + Forwardcandle
 - + Weekdaycandle
- ✘ Target vždy ternární
- ✘ Stoploss je součástí targetu
- ✘ TA Indikátory (některé) -> proměnné
- ✘ Nelze stavět na vizualizaci grafů (stroje nevidí..., zatím...)
- ✘ Je třeba normalizovat
 - + Rozsah -1;1
 - + Dělení směrodatnou odchylkou
 - + Využít TA indikátorů s předdefinovaným rozsahem (RSI)
- ✘ Nevizuální vícedimenzionální reprezentace (pro metodu nejbližšího souseda)
- ✘ Spojení indikátorů z více časových pásem
- ✘ Tvorba proměnných z více instrumentů

TYPY TARGETŮ



ONECANDLE

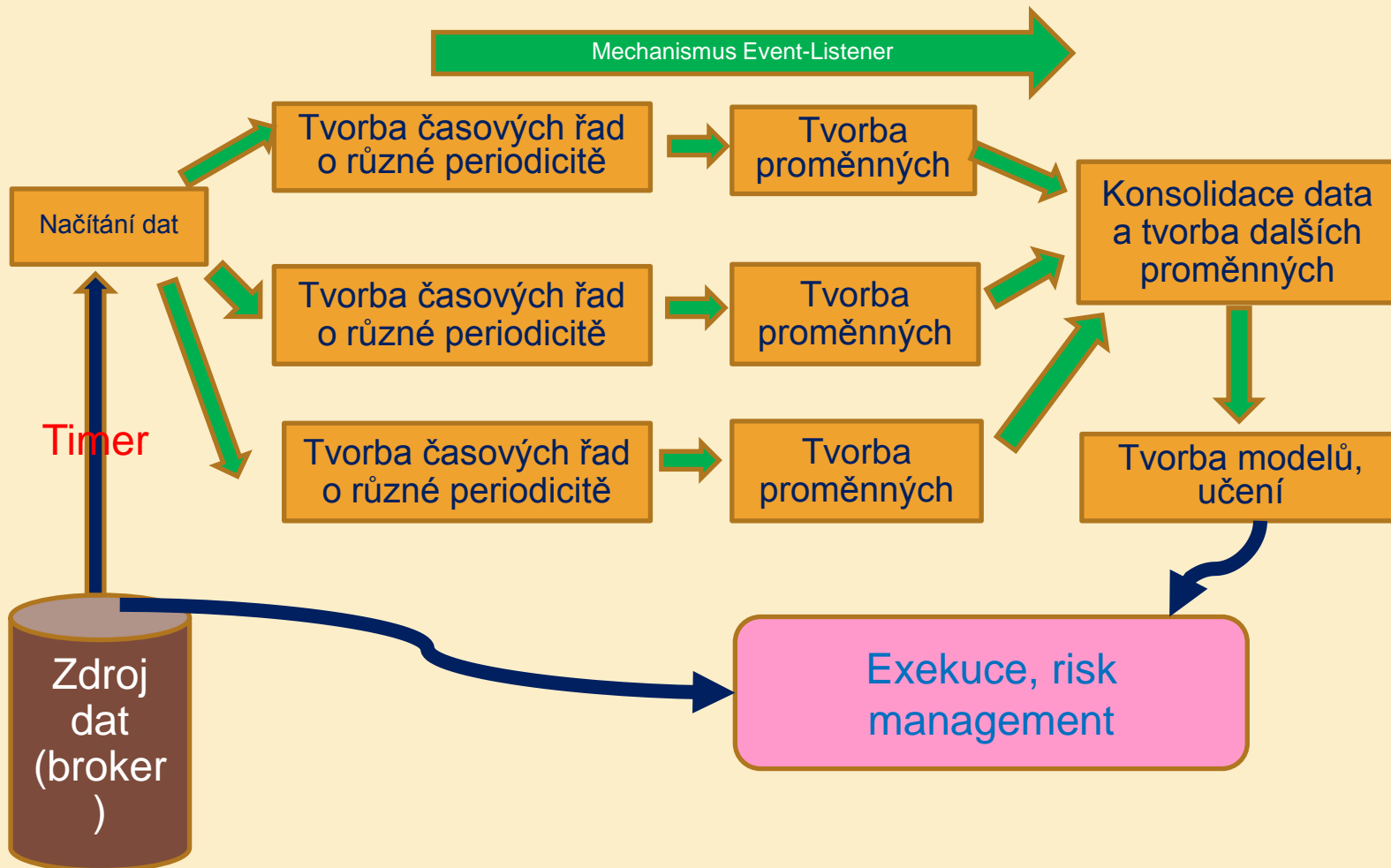
SOMECANDLE



PŘEHLED POUŽITÝCH MODELŮ

- × Testovala se:
 - + SVM
 - + Treebagger
 - + Lineární diskriminant
 - + Nearest-neighbor
 - + Distance Metric Learning
 - + Regrese (akcie), ne v Matlabu
- × Osvědčilo se
 - + Treebagger
 - + Nearest-neighbor

NÁVRH SYSTÉMU V MATLABU



NÁVRH SYSTÉMU V MATLABU

- × Použití Matlabu v celém procesu. Data se předávají mechanismem Event-Listener.
- × JE:
 - + Načítání dat
 - + Tvorba řad o různé periodicitě
 - + Tvorba proměnných (indikátorů) z dané časové řady
 - + Integrace řad, tvorba dodatečných proměnných založených od více instrumentů současně
 - + Učení modelu, evaluace modelu
- × Chybí
 - + Evaluace modelů online, test strategií online
 - + Napojení na brokera
 - + Risk management
- × Výhody:
 - + Jednotlivé kroky úlohy jsou oddělené
- × Nevýhody:
 - + Špatně se ladí chyby
 - + Kopírování velkých objemů dat mezi objekty
 - + Pomalé učení
 - + Velké nároky na paměť
 - + Velká pracnost

NĚKOLIK POZNÁMEK O MATLABU VE VZTAHU IMPLEMENTACI TRADING STRATEGIÍ

- ✘ Několik ukázek na webu Mathworks
- ✘ Není pro ultrarychlé obchodování, použitelný na několikasekundových timeframech a vyšších
- ✘ Otázka rychlosti a robustnosti prediktivních modelů
- ✘ Problematická spolupráce s některými proprietární programy (není to chyba Matlabu!)
- ✘ Výborná spolupráce s Javou
- ✘ API na Interactive Brokers (ActiveX, Java)

ZÁVĚR

- × Za pozitiva považujeme
 - + Definice targetu
 - + Implementace prediktivních modelů
 - + Transformace dat a tvorba proměnných
- × Nedporučujeme
 - + Používat mechanismus Event-Listener pro tento typ workflow
- × Dodělat je potřeba:
 - + Risk management, napojení na brokera, on-line testování

LITERATURA A ODKAZY

- × Api
- × www.undocumentedmatlab.com
- × <http://www.leptokurtosis.com/main/>
- × Backtesting
- × <http://wfatoolbox.com/>
- × Kvantitativní obchodování
- × <http://epchan.blogspot.com/>
- × www.erniechan.com
- × *Hutné čtivo (strojové učení, statistika)*
- × <http://vserver1.cscs.lsa.umich.edu/~crshalizi/notebooks/>
(Cosma Shalizi))
- × A samozřejmě www.mathworks.com