



Intelligentní forecasting

Roman Koch, Ondřej Bauer – 17. září 2020

Osnova



Motivace



Přehled metod – od historie po současnost



Ukázky a příklady



Budoucnost



Závěr a (diskuse ?)



jdeme na to ...

Motivace

- **Predikce je čím dál potřebnější pro rozhodování**
- **Množství časových řad je čím dál větší – např.:**
 - Mikrosegmentace portfolií
 - Větší množství sledovaných / měřených dat
- **Na predikce je čím dál méně času**
- **Zvětšuje se frekvence predikcí**
 - Běžně se dělají predikce na hodiny dopředu
 - Call centra
 - Řízení energetických soustav
 - Predikce spotřeby energie
- **Predikce se používají také pro:**
 - Detekce anomálií
 - Identifikace chyb v datech
 - Detekce podvodů

Motivace

- Velmi často se setkáváme s potřebou rychlé reakce na odchylky u mnoha řad najednou
- Není čas na „ladění“ super-parametrů jednotlivých metod / modelů
- Zvýšil se výpočetní výkon
- Zlepšili se metody postavené na bázi strojového učení (random forest, gradient boosting, ...)
- Existence hotových knihoven v Python se snadnou implementací
- Existence dostupných datových sad pro školení a testování robustnosti metod

Přehled metod

- Historické metody:
 - Průměr
 - Poslední hodnota
- Klasické metody:
 - Autoregression (AR) – 1920
 - Moving Average (MA) – 1901
 - Autoregressive Moving Average (ARMA) – 1970
 - Autoregressive Integrated Moving Average (ARIMA) – 1975
 - Seasonal Autoregressive Integrated Moving-Average (SARIMA) – 1976
 - Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors (SARIMAX)
 - Vector Autoregression (VAR)
 - Vector Autoregression Moving-Average (VARMA) – 1980
 - Vector Autoregression Moving-Average with Exogenous Regressors (VARMAX)
 - Simple Exponential Smoothing (SES) – 1960
 - Holt Winter's Exponential Smoothing (HWES)

Přehled metod - machine learning

- Prophet univariete, multivariete (by Facebook)
- Bayesian Ridge
- Lasso
- Random forest
- XGBoost – 2016
- Lightgbm (by Microsoft) – 2017
- SVM RBF (support-vector machines radial basis function)
- kNeighbors (Classifier implementing the k-nearest neighbors)
- Tensorflow simple LSTM (Long Short Term Memory) – 2015
- DeepAR (by Amazon)
- N-BEATS
- ESRRN
- ... a jejich kombinace
- a velmi rychle vznikají další

Co by nás mohlo překvapit

- Pro použití ML metod není potřeba extra dlouhé časové řady
- Tree models nepotřebují mnoho super-parametrů
- Pro ML modely je snadné přidávat externí (podpůrné) data (features)
- Není pravdou, že u ML modelů není možné vysvětlit závislosti

Budoucnost

Dnes se intenzivně pracuje na ML řešeních pro automatizovanou volbu nejvhodnější metody pro predikci v časových řadách a také pro správnou volbu super-parametrů – **Intelligent forecasting**

Mezi vhodné kandidáty je možné vzít např.: **autots**