

# Methodik BA

## 1. Laden von Datensets

- Dow Jones Index, [2013 - 2018]
- S&P 500 (Vaggel)
- (z.B. AAPL, BML, IBM, MSFT, PYP)

⇒ TECH Firmen

• Currency rates

[1985 - 2018] in US\$ (Vaggel)

• Goldpreis (Vaggel)

• Bitcoin-Kurs von [2014 - 2019]

• Coinbase (Vaggel)

(1-min data)

⇒ Anomalien (> 5% 1d ↑)

und Vorrat darlegen (Seaborn)

Idee: TECH stock

Datensets nehmen da

vergleichbar

## 2. Feature Extraction

- Features aus der technischen Analyse

(moving averages, momentum, ...)

da oft sehr sinnvoll als Eingangs-

zu "vorher OHLC"

• Feature Selection

bei Trees nicht nötig

da eingebettet (Entkopplung!)

↳ Importances

Idee (s. andere Paper):

Ich gehe von Best Trading

Strategie aus, z.B. Long/Short

über SD

## 3. Bestimmung Class Target

- Fehl. Fragestellung soll

z.B. sein: Soll ich diese Aktie kaufen, um in SD profitabel zu verhandeln?

⇒ 3A/NE/M (class.)

• Variation 1d/3d

oder Long-term

mit exponential smoothing

Idee (s. andere Paper):

Ich gehe von Best Trading

Strategie aus, z.B. Long/Short

über SD

## 4. Model Training & Testing

- 10-fold CV wird benutzt

→ DT, AdaBoost, RF + Baseline (Dummy Classifier)

Wdh. mit Hyperparameter (← Validation set)

(#-trees, max-depth, min-samples-split)

↓

Vergleich von Accuracy & F-Maß

(Exaktuell: Weighted RF (WQRTs), andere Abwandlungen)



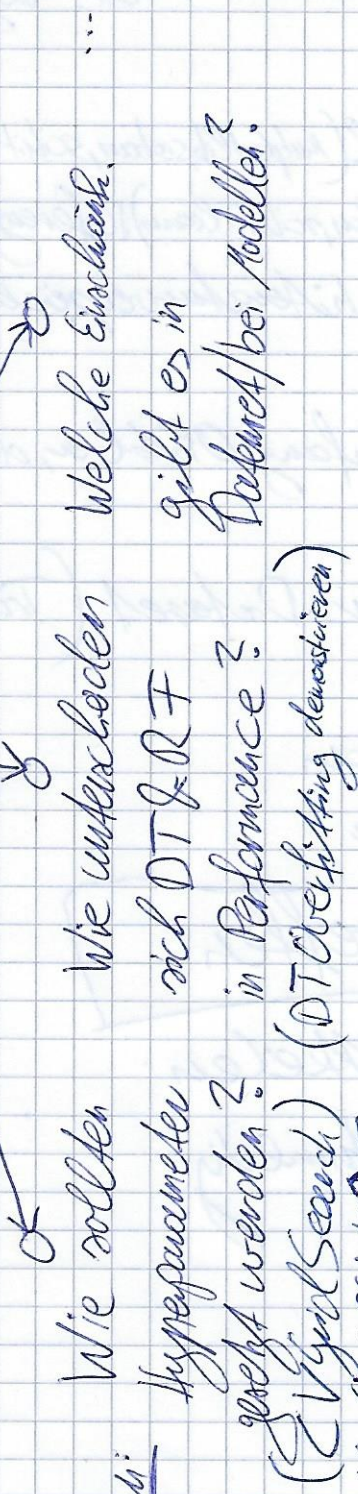
(Am Div. Richtig, geschätzt)

≙ gewinnbringend, Acc. > 50%

Klassifikation "Sollte immer hohe Genauigkeit in 50 Genauigkeit zu werden?"

invariant

FF: können DT & RFs (am Bsp. ~~Dataset~~) sinnvoll für Anforderungen eingesetzt werden?



Detaillfragen:

konfigurations-  
ansatz

Wie sollten Hyperparameter gesetzt werden? (Grid Search)   
  $\uparrow$   $\frac{\text{Ergebnis in Plot: } \text{acc} \uparrow}{\text{#trees}}$

Vergleich anhand Accuracy und Confusion Matrix (Recall, Precision, F1 Score) ist DT/RF auf diesem Dataset teilweise ungeeignet?

Fehler wichtige Features? Sind Hyperparameter nicht in Balance?

Dataset teilweise ungeeignet?

Ergebnis

Liste von Hyperparameter-Werten, die höchste Accuracy erzielen

Liste Einschränkungen