

# Analysing ensembles to improve forecast performance

## Objectives:

### 1. Analyse the performance of different ensemble techniques

Greater goal. Answered/achieved when answering the questions.

### 2. Specifically, analyse performance depending on different factors like

#### 2.1 The number of available forecasting models

N Zahl randomly untersuchen.

#### 2.2 the degree of similarity between available forecasting models

Ein Similarity Crterium herausfinden - Which are linear models - Which are ensembles models - which are glm

#### 2.3 the available training data

Falls genug Zeit dann Datenquelle berücksichtigen: UK DATEN.

### 3. Analyse how much do individual models contribute to an ensemble.

Can a model be a valuable contribution to an ensemble, even though it is not a good forecasting model if looked at in isolation?

#### 3.1 (General) Model drop

Random combinatorik. e.g. ein Ensemble mit 19 außer ein bestimmter.

Wie variable die Ensemble Score ist, wenn man modellen raus nimmst.

Auch abhängig von Modelgröße: e.g. 3 modellen rausholen ist relevanter für ein Ensemble von  $n=5$  als fuer  $n=20$ .

#### Example. How much do indivual models contribute in an ensemble of 20.

1. 20 nemhen. Baseline.
  - 10 mal (viel haufiger):
    - i. 2 randommly fallen lassen.

- ii. mit den 18 Score evaluieren.
2. Die 10 Scores dann
- i. Varianz berechnen.
  - ii. Haben sie sich sehr verschlechter in vergleich zu baseline?

Outcomes: - Kleine Varianz ist ein Indikator für Robustness. - wenn average verschlechterung: \$ Aussage (Bestätigung über NZahl.)

### **3.2 Specific model drop**

gezielt models rausholen: - Nach Qualität (die Schlechten raus) - oder nach einem anderen Kriterium.

Question: Inwieweit das Model improve, wenn man das Model rausnimmt oder reinnimmt.

### **4. Identify situations in which adding a model is beneficial or not**

- Nikos fragen.
- Heterogenität-Homogenität von Mitglieder (Bereits in 2.2 analysiert)

### **5. Analyse how stable forecast performance of different ensemble types is**

- Robustness/stability (are they the same?) already adresssed in 3.2