# Introduction (Selina)

* Introduction of the topic: what is the problem?
* Motivation: why the chosen topic is interesting and relevant?
* Goals: what is it the paper strives to achieve? What is the contribution?
* Outline: short presentation of the approach and structure of the paper KDD, erstes Protokoll, ausführlicher besprechen wie wir vorgehen/gedanken gemacht haben 1., 2. Iteration

# Preprocessing: Exploratory data analysis (felix/alex)

* (Explorative Datenauswertung/ Analyse (nicht für beide Iterationen), Statistik, Ziel variable verteilung etc.

# First iteration

## Preprocessing: Data cleaning

* (N/A values FREDDY und Outliers ALEX,

## Transformation

* Data reduction(correlation ALEX, FEATURE SELECTION SELINA, Encoding (scaling)MAX,

## Data mining

* (Anwendung der Algos, auswahl der algorithemn, Robust vs. Unrobost (2. Trainingssets) VERSTÄNDNISS AUFZEIGEN, HABEN GANZ VIEL GELERNT GASNZ TOLL FREDDY

## Evaluation

* (Maschinelle Sprachverarbeitung, argumentation, wo hackt es STARK ARGUMENTIEREN; SUPER GEIL) FREDDY/ MAX

# Second iteration

## Preprocessing: Data cleaning

* Ggf vorbereitung für feature selection, multidimensional outlier detection (nur 2 sätze was man machen könnte (mahalanobis distance, LOF), machen wir aber nicht, weil zu aufwendig) ALEX/MAX

## Transformation

* Data reduction(FEATURE SELECTION Selina, pca ALEX

## Data mining

* (cross validation, heterogene ENSAMBLE, Anwendung der Algos (Log regr, svm, random forest) incl feinere parameter selection, auswahl der algorithemn, Robust vs. Unrobost (2. Trainingssets) FREDDY

## Evaluation

* (Was hat sich verändert,) ALLE

# Conclusion

## Results

## Discussion

* Short summing up of the paper  What has been done?  What has been found?
* Discussion of key implications (take-aways)  Limitations/ ideas of future research

## Outlook