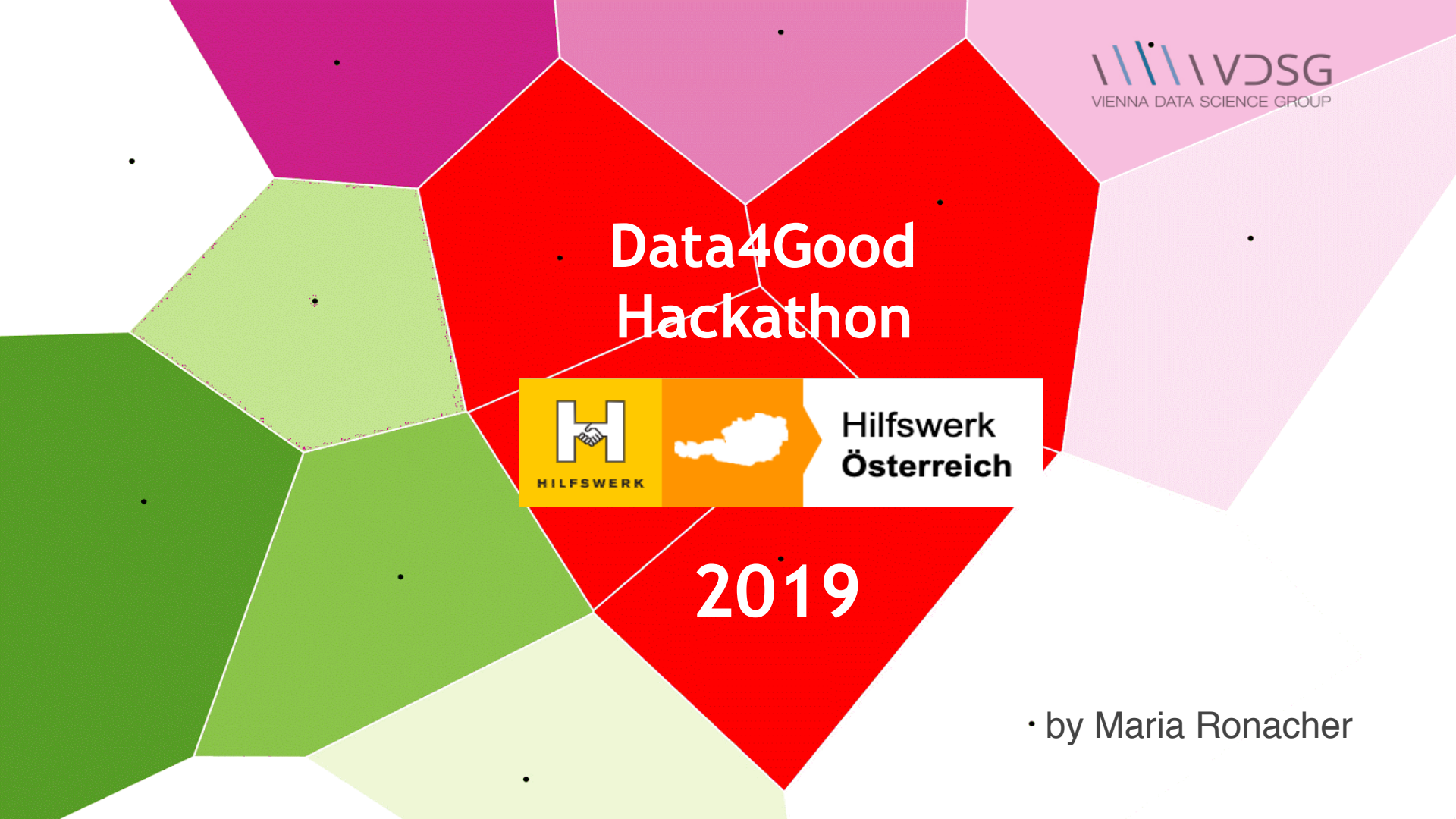

VDSG
VIENNA DATA SCIENCE GROUP
Data4Good
Hackathon
HILFSWERK
Hilfswerk
Österreich
2019
· by Maria Ronacher

Team

Project Lead

Maria Ronacher https://www.linkedin.com/in/mariaronacher/

Data Science Team

Bellinda Akwa-Asare https://www.linkedin.com/in/bellinda-akwa-asare-4538b52/

Jillian Augustine https://www.linkedin.com/in/jillianaugustine/

Matjaz Kovse https://www.linkedin.com/in/matjaz-kovse/

Roland Krasser https://www.linkedin.com/in/roland-krasser-5a2199139/

HILFSWERK

Hilfswerk Austria

24-hour care services

- Household services
- Support in lifestyle
- Keeping company
- Assistance with relocation
- Nursing and medical activities


CHALLEN
GE

Help revealing conflicts in ongoing relationships

Within the framework of their 24-hour care services for people with physical disabilities, Hilfswerk Österreich provide a matching service between patients and care providers. Cancellations are undesirable.

GOAL

Examine historical cases, in order to get insights on which data features are associated with cancellations.

24-Stunden-Betreuung - Visitenblatt

Betreute Person/en: _____ **RV/zust. Fachkraft:** _____ **Blatt Nr./Datum:** _____

Betreuer/in: _____

Beschäftigung	korrekt	mangelhaft	Tätigk. durch Angehörige	Anmerkungen/Änderungen zur Betreuungssituation/Maßnahmen
Gestaltung des Tagesablaufes	☐	☐	☐	
Einbeziehung KU in Tagesablauf	☐	☐	☐	
Freizeitgestaltung/Beschäftigung einfach/aufwändig	☐	☐	☐	
Haushaltsführung	korrekt	mangelhaft	Tätigk. durch Angehörige	Anmerkungen/Änderungen zur Betreuungssituation/Maßnahmen
Reinigung Haushalt	☐	☐	☐	
Reinigung persönl. Dinge (Prothese, Brille, Hörapparat, Leibstuhl),…	☐	☐	☐	
Sorgt für eine sichere und saubere Wohnumgebung, Beseitigung erkennbarer Gefahrenquellen für PB und KU	☐	☐	☐	
Wäscheversorgung	☐	☐	☐	
Versorgung der Tiere/Pflanzen	☐	☐	☐	
Müllentsorgung	☐	☐	☐	
Einkauf, Besorgungen und Botengänge	☐	☐	☐	
Essen und Trinken /Kochen	korrekt	mangelhaft	Tätigk. durch Angehörige	Anmerkungen/Änderungen zur Betreuungssituation/Maßnahmen
Unterstützung Nahrungs-und Flüssigkeitsaufnahme	☐	☐	☐	
Zubereiten Mahlzeiten	☐	☐	☐	
Adäquate Mahlzeit (Diät, Wunschkost,…)	☐	☐	☐	
Unterstützung Nahrungs- und Flüssigkeitsaufnahme bei Kau & Schluckstörungen	☐	☐	☐	

Records of visits and phone calls

The data about **cases** came as a set of PDF forms with records of visits and phone calls made by Hilfswerk employees.

- Two classes: normal or conflict
- Cases: few to several records in time order
- Records: provided services and conversations with the caregivers and patients

PREPARING THE DATA

1. Extraction
2. Anonymisation
3. Cleaning
4. Feature engineering
5. Splitting into training and test sets

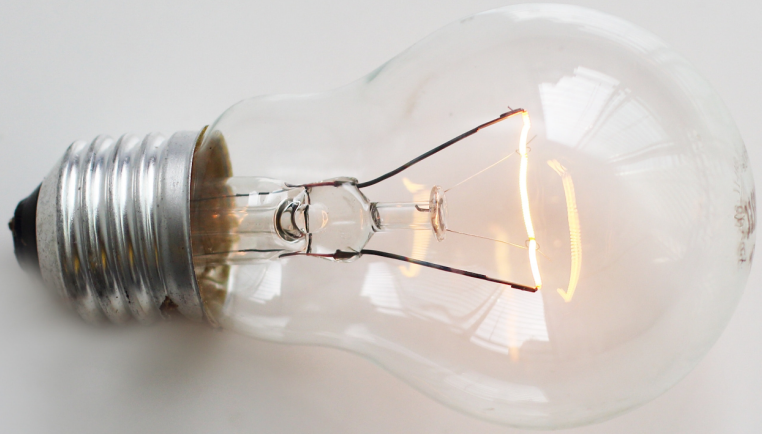
Hackathon objectives

- Text mining and sentiment analysis of the text data (conversations with the clients and caregivers)
- Time-series analysis of individual cases (conflict trends)
- Creating a machine learning model to predict status (normal/conflict)

Hackathon accomplishments

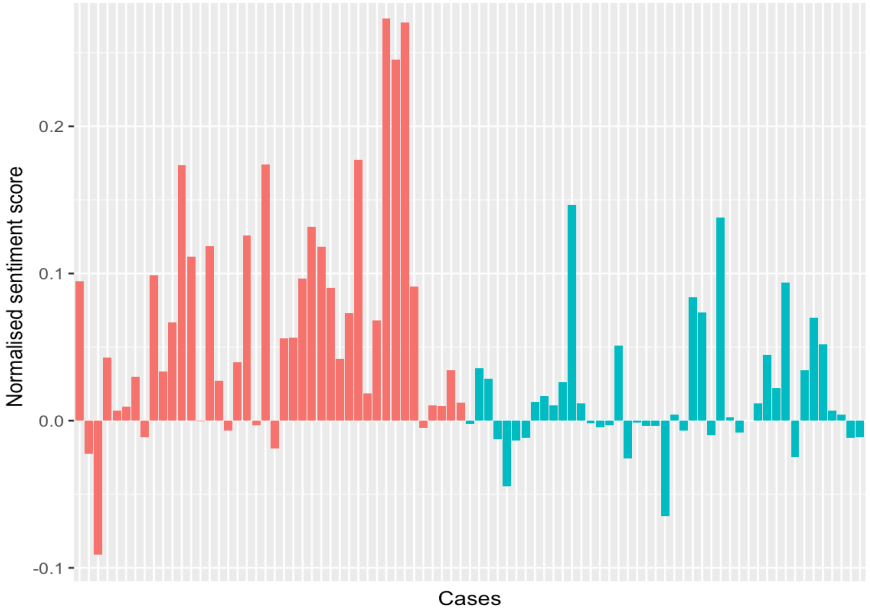
- Created multiple random forest classification models
- Performed text mining using different techniques
- Carried out sentiment analysis of conversations with clients and caregivers
- Engineer additional input variables which increased the performance of the models
- Assessed the feasibility of time-series analysis on the available data

INSIGHTS



Difference in sentiments

- Most normal and conflicted cases have positive sentiment in terms of German language, but the average sentiment score is lower for conflicted cases.
- Records of phone calls have more negative sentiments than records of visits. Larger numbers of telephone calls are associated with conflict.
- Sentiments calculated based on the dataset texts are the most important predictors of whether the case is a conflict or not (overfitting possible!)

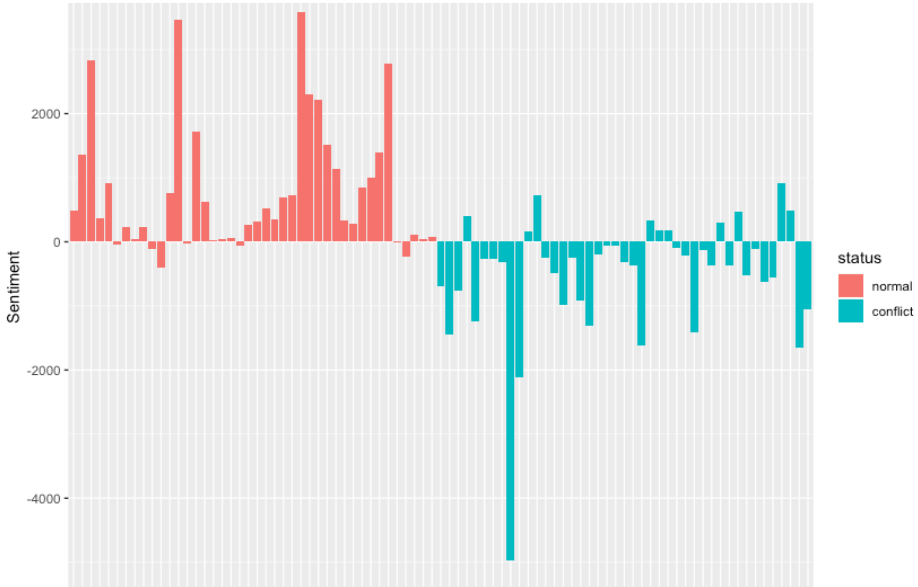


calculated based on sentiments of German words*

* SentiWS lexicon – a publicly available German-language resource for sentiment analysis

calculated based on the dataset texts*

* using tf-idf analysis of the normal and conflict conversations



Possible signs of conflict

- Frequent use of the word ‘nicht’ (especially by the caregiver)
- Usage of family words (‘Tochter’, ‘Mutter’, ‘Gattin’ etc.)
- Many conflicts happen in the first four months of care
- ...

INFERENCE RESTRAINT!

Need more data to validate the hypotheses.

POSITIVE

NEGATIVE

most frequent words

most prominent words
based on tf-idf analysis

POSITIVE

NEGATIVE

Conclusions

Although we had data and time limitations, we got some valuable results: findings, hypotheses and understanding how to deal with such challenges.

Road ahead

- Validate hypotheses on more data
- More thorough exploration of time trends in the data features, especially the conversations
- Further text mining, e.g. types of important words for normal/ conflict cases, caregivers/clients etc.
- Further improvement of the model, e.g. trying different feature sets

THANK YOU!

Website:
https://trendydots.com/data4good/challenge3.html

Maria Ronacher https://www.linkedin.com/in/mariaronacher/ maria@ronacher.eu

Questions?