

# Customer Feedback Sentiment Dictionary

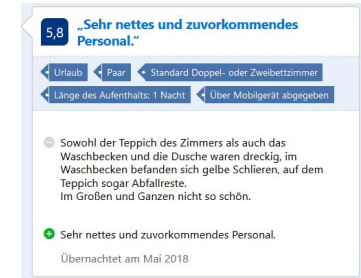
Towards Automatic Assessment of Online Hotel Reviews

## Idea

Guests assess and evaluate hotels constantly via digital channels.

Motivation

- Learning from this huge amount of data is crucial for future positioning in the market<sup>[1]</sup>
- Automatically created dictionary<sup>[2]</sup> identifies positive and negative connotated terms compared to their rating.



## Agenda

- Preparation of Study
- Data Analysis
- Results
- Conclusion

## Cornersto

Preparation



# Text - Preprocessing

- Standard steps<sup>[3]</sup>



- Further: separation of positive & negative reviews (via capitalization)
- Split data into a training corpus (80%) and validation corpus (20%)

Analysis

# Creation of the Dictionary using LASSO

Target criterion<sup>[4]</sup>:

$$RSS + \lambda \sum_{j=1}^p |\beta_j|$$

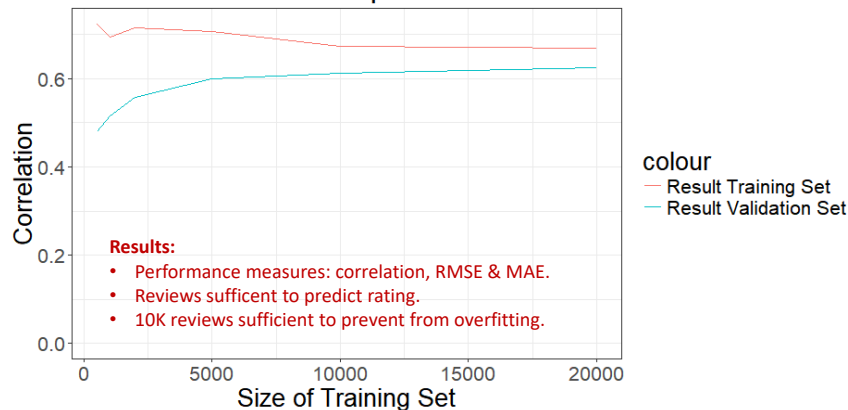
...equivalent to regression where  $\sum_{j=1}^p |\beta_j| < s$ .

- Term frequencies are used as predictors<sup>[2]</sup>.
- Regression coefficients  $\beta_j$  are shrunk towards zero.
- For LASSO some  $\beta_j = 0 \rightarrow$  implicit variable selection.
- Resulting coefficients  $\beta_j$  are weights of terms in resulting dictionary.

Analysis

# Predictive Power

Correlation between Expressions & Scores



**Results:**

- Performance measures: correlation, RMSE & MAE.
- Reviews sufficient to predict rating.
- 10K reviews sufficient to prevent from overfitting.

Result

# Resulting Sentiment Dictionary

| Positive     | Weight |
|--------------|--------|
| Phantastisch | 0.75   |
| ZUGEWIESEN   | 0.62   |
| genial       | 0.49   |
| traumhaft    | 0.49   |

| Negative               | Weight |
|------------------------|--------|
| RENOVIERUNGS-BEDÜRFTIG | -1.19  |
| DRECKIG                | -1.08  |
| JUGENDHERBERGE         | -1.03  |
| UNSAUBER               | -1.00  |

Size of the sentiment dictionary: 778 words.



Result

# Summary and Outlook

Conclusion

- ✓ A sentiment dictionary has been developed to assess positive & negative connoted words for hotel reviews.
- ✓ Written reviews are sufficient to quantify feedback in terms of rating scores.
- ✓ Investigation of size (#reviews) needed for a stable model.
- ✓ ...Future work may be dedicated to investigation of transfer possibilities.

# References

Conclusion

- [1] Rihova, I., Buhalis, D., Moital, M., & Gouthro, M.-B. (2015). Conceptualising customer-to-customer co-creation in socially dense tourism contexts. *International Journal of Tourism Research*, 17(4), 356-363. DOI: 10.1002/jtr.1993
- [2] Pröllochs, N., Feuerriegel, S. & Neumann, D. (2015). Generating Domain-Specific Dictionaries using Bayesian Learning. *ECIS 2015 Completed Research Papers*. Paper 144. DOI: 10.18151/7217444.
- [3] Feinerer, I., Kornik, K. & Meyer, D. (2008). Text Mining Infrastructure in R. *Journal of Statistical Software* 25(5). DOI: 10.18637/jss.v025.i05.
- [4] James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *Introduction to Statistical Learning*. Springer.

# Thank You

## Chancen der Digitalisierung nutzen

Optimale Lösungen für kleine und komplexe Projekte – wir unterstützen Sie auf dem Weg zur Digitalisierung Ihrer Geschäftsprozesse.

mehr erfahren

<https://www.kompetenzzentrum-rostock.digital/>