**Report Group INSERT\_GROUP\_NUMBER**

**Exercise 2**

WS 2021 - 188.977 Grundlagen des Information Retrieval

**Part1: Warmup**

TODO Insert your evaluation results into the following table

|  |  |  |
| --- | --- | --- |
| **Word1** | **Word2** | **Cosine Similarity** |
| cat | dog |  |
| cat | Vienna |  |
| Vienna | Austria |  |
| Austria | dog |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Word** | **Top-1** | **Top-2** | **Top-3** |
| Vienna |  |  |  |
| Austria |  |  |  |
| cat |  |  |  |

TODO: Analyze results briefly with a few words

**Part2: Short-Text Similarity**

TODO Insert your evaluation results into the following table

|  |  |  |
| --- | --- | --- |
| **Method** | **Preprocessing** | **Pearson Correlation** |
| Vector Space Model (from sklearn library) | Lowercasing+Stopword |  |
| Average Word Embedding | Lowercasing+Stopword |  |
| IDF Weighted Agg. Word Embedding | Lowercasing+Stopword |  |
| Vector Space Model (from sklearn library) | Lowercasing |  |
| Average Word Embedding | Lowercasing |  |
| IDF Weighted Agg. Word Embedding | Lowercasing |  |

TODO: Analyze results briefly with a few words.

**Part2: Training new language models**

|  |  |  |  |
| --- | --- | --- | --- |
| **Word (of your choice)** | **Top-1** | **Top-2** | **Top-3** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

TODO: Analyze results briefly with a few words.

TODO: Describe which data you used for training.

**Optional Section**

Here you can report:

* feedback
* found bugs/errors in the exercise or the lecture slides
* comments on your implementation that you want us to know
* interesting details/findings regarding the exercise, dataset, etc.
* ...