

Alberto Scinetti, 13115316  
Data Processing

## Process Book - Final Project

**Date:** 20/05

**What I've worked on:**

Learning about sentiment analysis with Vader to have a general understanding of this topic

**What problems I encountered:**

None

**What I learned:**

- How sentiment analysis works and what it is used for
- How Vader Sentiment Analysis is performed
- Vader polarity scores

**Which resources did I use:**

- > <https://medium.com/@rslavanyageetha/vader-a-comprehensive-guide-to-sentiment-analysis-in-python-c4f1868b0d2e>
- > <https://www.youtube.com/watch?v=QpzMWQvxXWk&t=1735s>

**Date:** 21/05

**What I've worked on:**

Working and learning about building a sentiment analysis tool with Python

**What problems I encountered:**

I tried to also use a already trained d library (RoBERTa Model for sentiment analysis) but I wasn't able to train the model successfully and implement it with example data (due to conflict of libraries that I wasn't able to resolve between PyTorch and tensor flow I believe)

**What I learned:**

- Further Topics on Sentiment Analysis with Python
- RoBERTa Model for sentiment analysis

**Which resources did I use:**

- > <https://www.youtube.com/watch?v=QpzMWQvxXWk&t=1735s>
- > [https://huggingface.co/docs/transformers/en/model\\_doc/roberta](https://huggingface.co/docs/transformers/en/model_doc/roberta)

**Date:** 22/05 - 24/05

**What I've worked on:**

Find the right dataset to perform my analysis on

Read the data, pre process and perform Vader Sentiment analysis  
Writing Code and the notebook

**What problems I encountered:**

/

**What I learned:**

- Further Topics on Sentiment Analysis with Python
- How to perform Vader Sentiment analysis
- Interpretation of Vader Sentiment Analysis and its polarity scores

**Which resources did I use:**

-> <https://medium.com/@rslavanyageetha/vader-a-comprehensive-guide-to-sentiment-analysis-in-python-c4f1868b0d2e>

-> <https://www.youtube.com/watch?v=QpzMWQvxXWk&t=1735s>

-> <https://www.geeksforgeeks.org/python-sentiment-analysis-using-vader/>

**Date:** 24/05

**What I've worked on:**

Topic modelling, understand how it's done and what results can be deduced  
Writing Code and the notebook

**What problems I encountered:**

/

**What I learned:**

- Topic modeling: what is it and how it can be performed
- Various approaches to topic modelling: LDA, Bert, top2vec

**Which resources did I use:**

-> <https://www.youtube.com/watch?v=bEaxKSQ4Av8>

-> <https://www.datacamp.com/tutorial/what-is-topic-modeling>

-> <https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0>

-> <https://doi.org/10.31046/tl.v11i1.506>.

**Date:** 25/05

**What I've worked on:**

LDA algorithm for topic modelling  
Writing Code and the notebook

**What problems I encountered:**

/

**What I learned:**

- How LDA algorithm is coded, works and implementation for topic modelling
- Interpret LDA words scores for topics and tag each topic correctly

**Which resources did I use:**

-> <https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0>

-> <https://bennett-holiday.medium.com/a-step-by-step-guide-to-writing-an-lda-program-in-python-690aa99119ea>  
-> <https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0>  
-> [https://www.youtube.com/watch?v=TKjilp5\\_r7o&ab\\_channel=PythonTutorialsforDigitalHumanities](https://www.youtube.com/watch?v=TKjilp5_r7o&ab_channel=PythonTutorialsforDigitalHumanities)

**Date:** 26/05

**What I've worked on:**

Top2Vec. Algorithm for topic modelling  
Writing Code and the notebook

**What problems I encountered:**

/

**What I learned:**

- How Top2vec algorithm is coded, works and implementation for topic modelling
- Interpret Top2vec words scores for topics and tag each topic correctly

**Which resources did I use:**

-> <https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0>  
-> [https://www.youtube.com/watch?v=TKjilp5\\_r7o&ab\\_channel=PythonTutorialsforDigitalHumanities](https://www.youtube.com/watch?v=TKjilp5_r7o&ab_channel=PythonTutorialsforDigitalHumanities)  
-> <https://medium.com/@janhavil1202/understanding-topic-modeling-with-top2vec-cdf58bcd6c09>  
-> <https://github.com/ddangelov/Top2Vec>

**Date:** 27/05 - 29/05

**What I've worked on:**

Refine the project, better visualization and write project description

**What problems I encountered:**

/

**What I learned:**

/