****

12/10/2022

**SUPERVISED LEARNING**

**Texte classification**

**Gedeon PASSO-KAFACK**

Table des matières

[Introduction 1](#_Toc121771000)

[Describe the dataset and task 1](#_Toc121771001)

[Proposed pipeline 1](#_Toc121771002)

[Experimental methodology 1](#_Toc121771003)

[Results and discussion 2](#_Toc121771004)

[Conclusion and next steps 2](#_Toc121771005)

[References 2](#_Toc121771006)

Introduction

# Describe the dataset and task

* + - What is the data about, who collected it and how?
    - What is the task targeted by your experiments?
  + **Summarize your objectives** and what you aim to achieve with your experiments

# Proposed pipeline

* + Briefly describe and **justify**the models you used and how you applied it/them to your task
    - Describe what your baseline model is and the more complex models you intend to try
  + Describe and **justify** any data pre-processing you had to do

# Experimental methodology

* + Describe and **justify** the methodology used to test your pipeline
    - Metrics you've evaluated
    - Cross-validation method (k-fold, single shuffle split, multiple shuffle splits, etc)
    - Hyper-parameter choice criteria (which were set to the framework default, which were tuned using grid/random/other search method)
  + Cite your GitHub repo for any implementation details

# Results and discussion

* + Present the **main relevant results** of your experiments
    - Use **tables and/or plots** to summarize the obtained performances on the train/valid/test sets
  + Based on your results, **explain your reasoning** behind changes you made to improve your performance
  + **Analyse and compare the results between baseline and proposed models**, discuss which is the best model for the task
  + Comment on the **main difficulties** encountered in running your experiments and/or to improve your models

Conclusion and next steps

* + Decide which model you would use for the task given your experimental results
  + Comment on eventual extra experiments or procedures you could perform to:
    - refine your decision,
    - to improve your model's performance or
    - to adapt your model to a particular application scenario (e.g. real-time inference, embedded inference, etc)

References

* + You do not need to explain every methodological step in detail if it is frequent practice in the field, but you should provide a reference in case the reader is not aware of it (a link to a documentation page, an article where the method is explained, or something of the sort)