Text Mining and Search: Medical abstracts

UNIMIB – Data science (2023/2024)

Group 10: Luca Milazzo (856702) e Marco Pagnini()

2024

Sommario

# Abstract

Intro+conclusions

# Introduction

OUR AIM + Introduction to the dataset + tools, explorative analysis

# Preprocessing

Preprocessing

# Text representation

Word2Vec, BERT, TF-IDF

# Classification

Solution + conclusions

# Clustering

Solution + conclusions

# Conclusions

General conclusions + future extensions

# Bibliography