

Project

# Evaluation of different classification algorithms for COVID-19 Pandemic Tweets

Text Mining (732A92)

Stylianos Sidiropoulos (stysi607)



Division of Statistics and Machine Learning  
Department of Computer Science  
Linköping University

01-2021

# Contents

<b>1</b>	<b>Abstract</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>3</b>
<b>3</b>	<b>Data</b>	<b>4</b>
3.1	Data Preprocessing . . . . .	4
3.2	Natural Language Processing (NLP) . . . . .	4
<b>4</b>	<b>Statistics</b>	<b>5</b>
4.1	Number of words in tweet . . . . .	5
4.2	Wordclouds . . . . .	5
4.3	Uni-grams . . . . .	5
4.4	Bi-grams . . . . .	5
<b>5</b>	<b>Classification methods</b>	<b>5</b>
5.1	Multinomial Naive Bayes . . . . .	5
5.2	Logistic Regression . . . . .	5
5.3	Random Forest Classifier . . . . .	5
5.4	Decision Tree Classifier . . . . .	5
5.5	Linear Support Vector Classifier (SVC) . . . . .	5
5.6	Stochastic Gradient Descent (SGD) Classifier . . . . .	5
5.7	Extreme Gradient Boosting (XGB) Classifier . . . . .	5
5.8	Neural Networks . . . . .	5
<b>6</b>	<b>References</b>	<b>5</b>

# 1 Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis auctor efficitur nisi ac ultricies. Nulla et viverra diam, ut interdum nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nulla facilisi. Aenean consequat aliquet elit eu semper. Fusce posuere quis purus sit amet pretium. Etiam ultrices finibus neque, quis laoreet justo sagittis in.

Etiam iaculis ultricies imperdiet. Cras purus nisl, vehicula nec ipsum hendrerit, tristique bibendum sem. Aenean sed urna euismod, volutpat augue aliquet, congue turpis. Mauris porttitor tempor volutpat. Aliquam ac dui lectus. Maecenas venenatis tempor nibh eu volutpat. Mauris quis posuere est. Nulla purus nisl, viverra facilisis ipsum id, feugiat lacinia lectus. Etiam sodales fermentum euismod. Nulla facilisi. Ut laoreet gravida nisi, at imperdiet magna vestibulum nec. Fusce rutrum ante magna, ut pulvinar sapien mollis posuere.

## 2 Introduction

Integer pulvinar nulla ac quam ultricies, a ornare lorem sagittis. Morbi quis nisl vulputate, maximus mauris vel, tincidunt ipsum. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec id nisl vehicula, vulputate magna ut, placerat eros. Fusce in mauris ullamcorper, imperdiet sapien at, interdum dolor. Nam turpis erat, ornare in tristique sit amet, convallis ut dolor. Nam id odio interdum, rutrum lacus sit amet, dictum nisl. Phasellus aliquet vulputate eros id euismod. Vivamus orci urna, faucibus tincidunt nisi vel, semper vehicula leo. Curabitur tristique molestie semper. Sed dictum auctor efficitur. Quisque et augue bibendum, tincidunt magna vel, posuere nisl. Pellentesque quis placerat purus. Vestibulum sed sem est. Pellentesque sit amet porttitor sapien. Integer at enim eget ligula tempor cursus vitae nec purus.

Nulla velit tellus, tincidunt quis cursus id, tristique sit amet eros. Fusce cursus est vitae mi dapibus pellentesque. Sed elementum est et ante convallis placerat. Nullam consectetur, felis laoreet interdum tristique, mauris felis congue dolor, ac fermentum orci leo non dolor. Fusce varius ligula sit amet urna ultricies ornare. Sed placerat elit sit amet cursus consectetur. Maecenas non nisi ligula. Ut venenatis nisl ipsum, vitae pellentesque mauris consectetur eget. Pellentesque vestibulum velit ut ultrices ullamcorper.

Fusce non nisi mi. Donec maximus, arcu quis tempus consequat, leo augue imperdiet velit, et interdum arcu dui non quam. Nullam vel neque sed nulla blandit porta sed id quam. Aliquam id sem urna. Morbi pellentesque lacinia pharetra. Integer tincidunt massa convallis, ullamcorper metus ac, vulputate turpis. Curabitur dictum lectus in sem blandit, sed rutrum risus tincidunt. Curabitur posuere sapien sit amet enim faucibus, vel maximus justo lobortis. Sed sollicitudin dictum rhoncus. Aliquam quis gravida dui, eu consectetur elit. Cras interdum erat ipsum, at sagittis enim accumsan vitae. Pellentesque id iaculis sapien, sit amet ultrices lectus. Fusce tempor est justo, nec laoreet dui consectetur porttitor. Vivamus consectetur tempus mattis. Vestibulum ornare tincidunt consectetur.

## 3 Data

Mauris vestibulum venenatis turpis eu blandit. Sed est tellus, auctor id tortor a, pellentesque consequat quam. Phasellus efficitur elit et rutrum egestas. Phasellus sagittis leo non sollicitudin tincidunt. Vivamus rhoncus iaculis aliquam. Proin pretium, risus ut tempor consectetur, nunc augue pretium est, at cursus velit justo vel arcu. Proin dictum ligula a commodo rutrum.

### 3.1 Data Preprocessing

Vestibulum non eros tortor. Nullam arcu ante, placerat non rutrum vitae, blandit eget elit. Suspendisse potenti. Interdum et malesuada fames ac ante ipsum primis in faucibus. Sed placerat euismod est nec vulputate. Ut id metus vitae purus pharetra placerat et et sapien. Maecenas lacinia scelerisque leo nec facilisis. Etiam ut mollis dolor. Sed id bibendum elit, non ullamcorper nisl. Morbi sodales mollis commodo. Morbi cursus odio ut mattis accumsan. Aenean ac felis dignissim, luctus est sed, pulvinar velit. Integer turpis elit, posuere nec erat ac, hendrerit tristique ligula.

### 3.2 Natural Language Processing (NLP)

Phasellus ac consequat nibh, vel condimentum quam. Morbi in est ut odio pharetra feugiat a quis nulla. Donec rhoncus eget nunc vitae ullamcorper. Sed quis bibendum ex. Nullam non augue nec leo facilisis facilisis ullamcorper non turpis. Nullam rhoncus, ipsum non efficitur viverra, velit ligula hendrerit felis, porta feugiat eros urna ac justo. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia curae; Nam gravida consectetur rhoncus. Vivamus suscipit ut nulla non vestibulum. Phasellus maximus nisl non odio porta, vel volutpat ligula volutpat.

Vestibulum eu sapien ante. Maecenas lorem libero, tincidunt ut accumsan eget, sodales vitae diam. Pellentesque lobortis aliquet ligula, et condimentum tortor finibus a. Etiam ullamcorper enim eget sem fringilla, at viverra ante tincidunt. Fusce diam felis, cursus in urna tincidunt, accumsan pretium est. Duis vulputate tellus a ultrices varius. Integer accumsan orci ac risus dignissim eleifend. Ut hendrerit ligula ac tellus luctus molestie. Pellentesque consequat egestas leo, vitae vulputate tellus viverra ac. Nunc tincidunt fermentum sagittis. Integer mi erat, rhoncus at sem eu, dapibus malesuada leo. Mauris diam nisi, sagittis eu fermentum eu, venenatis non tortor. Suspendisse et viverra felis, non posuere nisi. Praesent congue libero quis augue imperdiet euismod.

## 4 Statistics

### 4.1 Number of words in tweet

### 4.2 Wordclouds

### 4.3 Uni-grams

### 4.4 Bi-grams

## 5 Classification methods

### 5.1 Multinomial Naive Bayes

### 5.2 Logistic Regression

### 5.3 Random Forest Classifier

### 5.4 Decision Tree Classifier

### 5.5 Linear Support Vector Classifier (SVC)

### 5.6 Stochastic Gradient Descent (SGD) Classifier

### 5.7 Extreme Gradient Boosting (XGB) Classifier

### 5.8 Neural Networks

## 6 References