KLASIFIKASI KATEGORI BERITA MENGGUNAKAN ALGORITMA SUPPORT VERTOR MACHINE

Aaz Muhammad Hafidz Azis1, Khansa2, dst….3,

Teknik Informatika, Fakultas Sains dan Teknologi

Universitas Islam Negeri Sunan Gunung Djati Bandung

11770500011; 11770500322; 1177050037 3

ABSTRAK

## Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sagittis sem pellentesque laoreet commodo. Maecenas nec diam orci. Morbi tristique mi a leo porta lacinia. Vivamus rutrum urna eget nunc sollicitudin, quis molestie urna volutpat. Proin bibendum consequat justo id suscipit. Pellentesque fermentum ligula vitae felis suscipit ullamcorper. Integer varius in urna fringilla sollicitudin. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Sed ultricies ut urna id vehicula. Nam varius orci aliquam quam aliquet, ac efficitur arcu vulputate..

## Kata Kunci : Lorem ipsum dolor sit amet

# PENDAHULUAN

## Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sagittis sem pellentesque laoreet commodo. Maecenas nec diam orci. Morbi tristique mi a leo porta lacinia. Vivamus rutrum urna eget nunc sollicitudin, quis molestie urna volutpat. Proin bibendum consequat justo id suscipit. Pellentesque fermentum ligula vitae felis suscipit ullamcorper. Integer varius in urna fringilla sollicitudin. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Sed ultricies ut urna id vehicula. Nam varius orci aliquam quam aliquet, ac efficitur arcu vulputate.

## Vestibulum at facilisis risus, quis varius arcu. Mauris eu pellentesque libero. Nam volutpat, augue et feugiat vestibulum, lacus libero fringilla tortor, ullamcorper porta nunc lorem vel ante. Donec hendrerit ullamcorper lectus id posuere. Nulla vitae porttitor tortor. Maecenas iaculis metus non malesuada viverra. Duis dapibus lacus ac nisi feugiat, sed auctor felis placerat. Nullam lobortis efficitur iaculis. Suspendisse augue mi, consequat a libero ac, feugiat commodo mi. Cras massa purus, fermentum a ligula nec, tincidunt euismod ligula. Maecenas blandit aliquet sapien quis sagittis. Maecenas eget leo in sapien tempor ultricies. Etiam ac enim et lectus vestibulum posuere ut ac turpis. Nullam sapien urna, egestas quis arcu vitae, elementum pharetra nisi.

## Sed finibus dui eros, quis sodales mi gravida ut. Integer porta lorem ac nibh egestas posuere. Nunc fringilla, diam nec feugiat lobortis, nisl est convallis felis, vel sollicitudin risus elit ut dolor. Integer tempus, lorem eget cursus semper, est elit pulvinar ante, eget porttitor elit felis et sapien. Pellentesque nibh ipsum, vulputate nec mi sed, interdum luctus dui. Donec velit ex, rhoncus ac blandit eu, congue in sem. Ut in rutrum mi. Sed tellus eros, iaculis eu lorem sed, blandit sagittis quam. Nulla non fermentum erat. Quisque tincidunt tempus nulla, eget fringilla odio. In non mauris in nisl cursus luctus eget sed est. Duis in augue a mauris fringilla viverra. Nulla eu commodo sapien. Curabitur rutrum justo nec leo fermentum condimentum.

## Maecenas auctor sollicitudin enim sit amet tristique. Pellentesque consequat orci ex, id facilisis tortor ultrices nec. Pellentesque ultricies nisl vitae lacus tristique venenatis. Vestibulum vestibulum sodales scelerisque. Mauris ut faucibus augue. Proin eu faucibus orci, vel porta urna. Nam cursus lectus aliquam feugiat mattis. Duis quam odio, fermentum vel semper sed, rutrum nec sem. Sed elementum luctus arcu. Curabitur augue ante, varius nec tempor ac, congue ac nisl. Nam nisl enim, porta vitae commodo et, ultricies pellentesque tellus. Vivamus tristique sapien vel semper accumsan.

## Vestibulum placerat ultricies neque quis tempor. Ut et eleifend libero, eget vulputate nunc. Praesent sit amet eros ex. Donec auctor, nunc quis pretium tincidunt, leo lectus aliquam ante, non rutrum tortor diam ac dui. Quisque justo mauris, pretium condimentum tristique et, fringilla sit amet nunc. Nulla magna turpis, tincidunt a dictum nec, luctus eu risus. Sed consectetur nunc eget dignissim dignissim. Vestibulum a dolor diam. Integer maximus mattis velit. Nunc nec consectetur nisl. Nullam eu purus ac metus suscipit egestas nec sit amet erat. Vivamus laoreet blandit tortor, scelerisque efficitur libero fermentum sed. Pellentesque facilisis ex ac pellentesque ullamcorper..

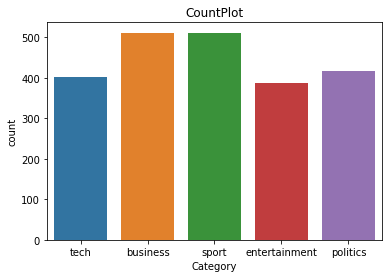
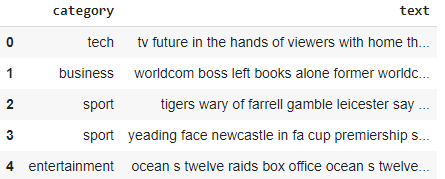
# METODE STUDI

## Di bagian ini, kami mendeskripsikan mengenai pengaplikasian atau implementasi algoritma Support Vector Machine (SVM) dalam mengklafisikasikan berita berdasarkan kategorinya. . Algoritma SVM melakukan klasifikasi secara *supervised learning* atau mempunyai input dan output yang telah dibentuk menjadi suatu model hubungan matematis yang dapat melakukan klasifikasi dan prediksi dari data yang sudah ada sebelumnya.

## Adapun tahapan yang dijelaskan seperti berikut.

## *Gambar 1. Alur klasifikasi berita*

## Dalam proses pengumpulan data, dataset yang digunakan berasa dari kumpulan berita BBC News yang telah diberikan kategori atau label. Label ini terdiri atas Teknologi, bisnis, sport, entertaimen, dan politik. Yang kemudian dataset ini digunakan untuk training terhadap system agar dapat memprediksi berita-berita lainnya.



## *Gambar 2. Kumpulan Data di seluruh negara*

## Tahapan preprocessing terhadap dataset ini merupakan proses pemilahan data menjadi data yang bias langsung di lakukan training menggunakan algoritma svm, dalam melakukan proses ini terbagi menjadi 5 tahapan yaitu

## 1. Removing blankspace.yaitu penghapusan spasi yang terdatap dalam kalimat sehingga menyisakan kata-kata yang diperlukan untuk data latih. .

## 2. Casefolding, yaitu pengubahan penulisan menjjadi lowcase pada semua kata/huruf.

## 3. Tokenizing, yaitu pemotongan pada setiap kalimat menjadi beberapa bagian. Setiap kalimat tersebut dikumpulkan dalam bentuk list

## 4. Encoding. Yaitu tahapan yang merubah kata-kata atau huruf menjadi bentuk vector atau numerik

## Pada tahapan klasifikasi menggunakan SVM ini, data dibagi menjadi dua bagian yaitu data train dan data test yang nantinya digunakan sebagai evaluasi dari kinerja SVM dalam melakukan proses klasifikasi. Data latih yang digunakan pada tahapan ini sudah dilakukan melewati tahap processing yang sudah siap untuk dilakukan training.

## Tahapan akhir pada proses ini yaitu evaluasi yang mengukur tingkat akurasi, f1 scre dari proses klasifikasi. tahapan-tahapan preprosecing tampilkan dalam gambar 2.

# HASIL DAN PEMBAHASAN

## Dalam penelitian ini ,kami menggunakan data yang merupakan data sekunder yang berupa data berita dari BBC News. Pengolahan data ini diimplementasikan menggunakan Bahasa python perbandingan data pada tahapan preprosesing dan sebelum dilakukan preprocessing dapat dilihat pada gambar 2.

## 

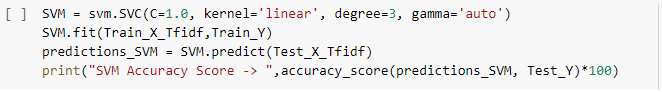
## *Gambar 3. Kumpulan Data di Benua Asia*

Tapahap encoding ini merubah bentuk teks menjadi numerik. Hal ini silakukan sebelum proses training menggunakan algoritma SVM. tampilan data setelah dilakukan proses encoding ditunjukkan pada Gambar 3.

## 

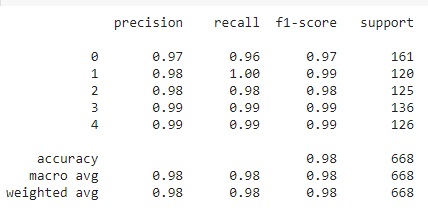
## *Gambar 3. Proses Training*

Implementasi Algoritma SVM ini menggunakan library yang sudah disediakan dengan source code yang ditampilkan pada gambar .



## *Gambar 3. Proses Training*

Hasil klasifikasi diperoleh nilai akurasi mencapai 98.35 % dengan rata rata presisi 90%, recall 98%, F1-Score 98% dan Support sebesar 668. Hasil implementasi ditampilkan pada gambar



. *Gambar 4. Training Loss dan Accuracy COVID-19 dataset*

# KESIMPULAN

## Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sagittis sem pellentesque laoreet commodo. Maecenas nec diam orci. Morbi tristique mi a leo porta lacinia. Vivamus rutrum urna eget nunc sollicitudin, quis molestie urna volutpat. Proin bibendum consequat justo id suscipit. Pellentesque fermentum ligula vitae felis suscipit ullamcorper. Integer varius in urna fringilla sollicitudin. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Sed ultricies ut urna id vehicula. Nam varius orci aliquam quam aliquet, ac efficitur arcu vulputate.

# DAFTAR PUSTAKA

[1]        S. P. Adhikari *et al.*, “Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: A scoping review,” *Infectious Diseases of Poverty*. 2020, doi: 10.1186/s40249-020-00646-x.

[2]        “Virus Corona (COVID-19) - Gejala, penyebab dan mengobati - Alodokter.” https://www.alodokter.com/virus-corona (accessed May 08, 2020).

[3]        U. Of, U. C. As, D. For, S. Protection, T. O. Prevent, and C. Virus, “PEMANFAATAN UV-C CHAMBER SEBAGAI DISINFEKTAN ALAT PELINDUNG DIRI UNTUK PENCEGAHAN PENYEBARAN VIRUS CORONA UTILIZATION OF UV-C CHAMBER AS DISINFECTANT FOR SELF PROTECTION TOOL,” vol. 7, no. April, 2020.

[4]        J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, “COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection,” 2020, [Online]. Available: http://arxiv.org/abs/2003.12338.

[5]        T. Ai *et al.*, “Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases,” *Radiology*, 2020, doi: 10.1148/radiol.2020200642.

[6]        A. Narin, C. Kaya, and Z. Pamuk, “Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks,” 2020, [Online]. Available: http://arxiv.org/abs/2003.10849.

[7]        A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” pp. 1–9, 2020, [Online]. Available: http://arxiv.org/abs/2003.13815.

[8]        F. Shi *et al.*, “Large-Scale Screening of COVID-19 from Community Acquired Pneumonia using Infection Size-Aware Classification,” 2020, [Online]. Available: http://arxiv.org/abs/2003.09860.

[9]        Y. Song *et al.*, “Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images,” *medRxiv*, 2020, doi: 10.1101/2020.02.23.20026930.

[10]       E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, “COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images,” 2020, [Online]. Available: http://arxiv.org/abs/2003.11055.