**­­­­­Project Mandiri NLP (Final Report)**

Nama: Muhammad Rizqi

NIM: 0110221267

**Topik**

Resume Screening

**Referensi:**

* Resume Screening using Machine Learning: <https://www.kaggle.com/code/gauravduttakiit/resume-screening-using-machine-learning>
* Resume-Screening-with-NLP: <https://www.kaggle.com/code/akashkotal/resume-screening-with-nlp>
* Resume Screening App: <https://github.com/611noorsaeed/Resume-Screening-App>

**Link repository project:** [**https://github.com/Rizqi-25/ProjectNLP/tree/master**](https://github.com/Rizqi-25/ProjectNLP/tree/master)

1. **Pembuatan Model**

A screenshot of a computer

Description automatically generated

Terdapat 3 model yang digunakan yaitu KNN, SVC, dan RandomForestClassifier.

1. KNN

Proses training model KNN.

A screenshot of a computer

Description automatically generated

1. Model SVC

Proses training model SVC.

A screenshot of a computer program

Description automatically generated

1. Model RandomForestClassifier

Traning model RandomForestClassifier.

A screenshot of a computer

Description automatically generated

Menyimpan model dengan pickle.

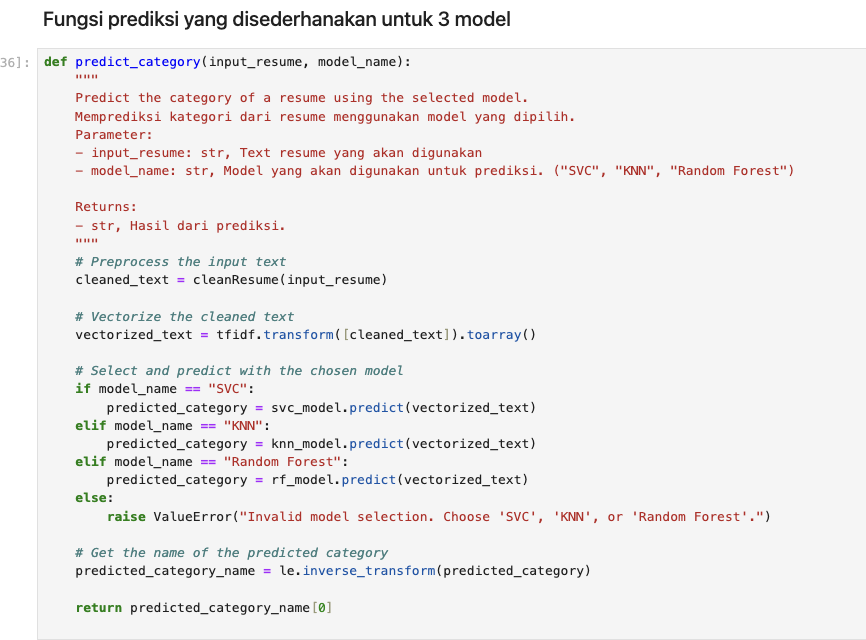
A screenshot of a computer code

Description automatically generated

1. **Demo Hasil Project**

Setelah semua model dibuat, Langkah selanjutnya membuat fungsi prediksi resume untuk mengetest model-model yang telah dibuat.

1. **Fungsi Prediksi**

****

1. **Testing**

**A screenshot of a computer

Description automatically generated**

1. **Demo Menggunakan Streamlit**

Dibuat juga webapp untuk prediksi CV menggunakan Streamlit.

A screenshot of a computer

Description automatically generated

Pada webapp streamlit kita dapat memilih model apa yang ingin digunakan untuk memprediksi CV. CV dapat diupload dengan format PDF, TXT, ataupun DOCX.

1. **Proses Prediksi pada Streamlit**

Setelah memilih model yang diinginkan dan mengupload file cv, berikut hasil dari prediksinya.

A screenshot of a computer

Description automatically generated

Ditampilkan juga pada webapp hasil ekstrasi text pada cv yang telah melalui proses TF/IDF.

1. **Kesimpulan**

Dari ketiga model yang dibuat, ketiga model tersebut memiliki tingkat akurasi yang tinggi dengan KNN sebesar 98.45%, SVC 99.48%, dan RandomForestClassifier 99.48%. Hasil test juga menunjukan prediksi yang diberikan sesuai dengan kriteria pada CV.