**INPUT:**

- User's answers to the questionnaire on the website

- CV in PDF format

**OUTPUT:**

- List of Top Predicted Jobs sorted

- List of Top features

- A recommended route for finding the desired job

**DATA SOURCES:**

Resumes from Kaggles sorted by categories.

**DATA PROCESSING:**

projectest file - Conversion of PDF files to DF characterized by categories using functions.

**RECOMMENDED ALGORITHMS:**

Transformer Models like BERT / GPT / T5 :

These models can be used to match the subject descriptions to the education program descriptions in a more accurate way.

Sequence Models like Recurrent Neural Networks (RNN) / Long Short-Term Memory (LSTM) / Gated Recurrent Units (GRU):

These models are suitable for analyzing serial or textual data and can deal with the description of the profession and education programs in a good way.

LSTM or GRU can be used to analyze textual descriptions of the subjects and educational programs and find the correspondence between them.

**RISKS:**

Transformer Models (BERT / GPT / T5):

**High Resource Consumption**: These models require significant computational power and memory, leading to high operational costs.

**Complex Implementation**: Implementing and maintaining these models requires a high level of expertise, which can be a barrier for smaller teams or those with limited technical knowledge.

**Ethical Issues and Bias**: These models may include biases if trained on unbalanced data, potentially resulting in unfair outcomes

Sequence Models (RNN / LSTM / GRU):

**Model Design Challenges**: Designing these models correctly can be challenging, especially for very long sequential data.

**Long-Term Dependencies**: These models can struggle with handling long-term dependencies in the text.

**Processing Time**: They may be slow in systems with limited resources, especially for long texts.

**TIMES:**

Until the end of July.