Resume Parsing Challenges

1. Reading Documents

- Converting any document into text
- Reading table format resume (single page multiple column)
- Reading data from tables

2. Extraction of data

- Setting rules for every possible format
- Extracting data with special characters and white spaces
- Defining section limits for extracting particular data
- Identify and map the data extracted from table

3. labeling of Data

- Labeling ambiguous Data
- Creating Vectors without losing important keywords
- Summarizing data

4. Identification of synonyms

- Finding Source of Synonyms
- Storage strategy (Whether to store the data or use API)

- Efficient algorithm to identify and replace synonyms with actual word in linear time complexity