wowlabz CASE STUDY

Resume Parser

What can be the quickest method for presenting a relevant group of candidates to a potential recruiter?

CLIENT
Resume Parser

PRODUCT
Resume Parser

PLATFORM
Web App

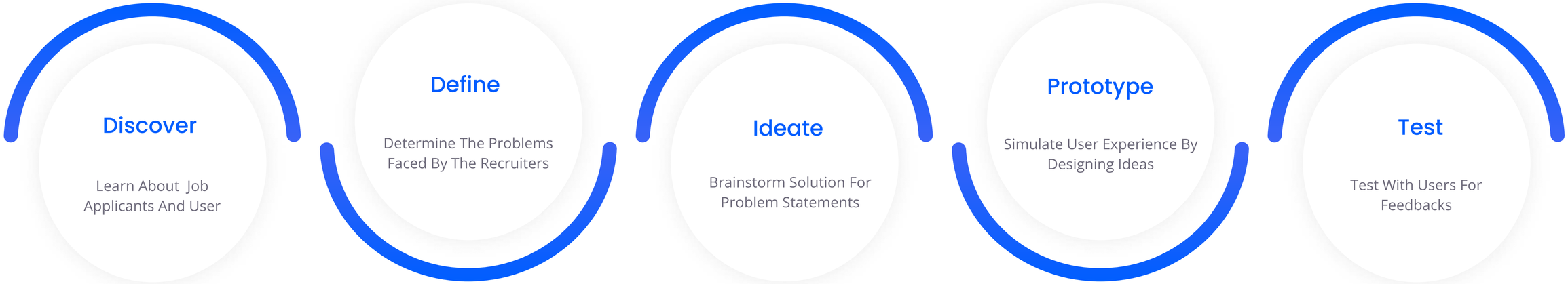
TIMELINE
10 weeks

DESIGNERS
1 Design Lead
1 UX/UI Designer

Overview

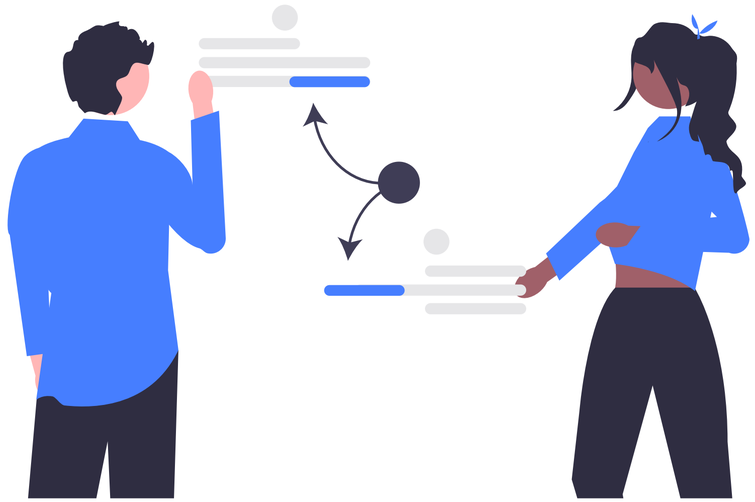
OBJECTIVE

The aim of the project was to create a tool for use in the recruitment industry that accelerates the recruitment process by extracting candidate information from unstructured resumes and presenting it in a tabular format. The following steps were taken to achieve the objective of this project:



Problem Statement

- Identifying a pool of qualified candidates, screening them for the desired skills and experience, and then selecting the most suitable candidate for the job.
- Every candidate has a resume in different formats and structures (an unstructured data type).
- For employers and recruiters, it is difficult to identify the skills and qualifications of potential candidates in unstructured resumes.
- Nothing in structured tabular format.



Possible **Solution**

- The problem of parsing the unstructured document can be solved using deep learning methods like **Name Entity Recognition(NER), and Object detection.**

What Is Object Detection ?

Object detection is the task of detecting instances of objects of a certain class within an image. The state-of-the-art methods can be categorized into two main types:

1. One-stage methods
2. Two-stage methods

One-stage methods prioritize inference speed, and example models include YOLO, SSD and RetinaNet.

Two-stage methods prioritize detection accuracy, and example models include Faster R-CNN, Mask R-CNN and Cascade R-CNN. Object detection algorithms can detect the different sections of a resume like experience, education, personal details, etc.

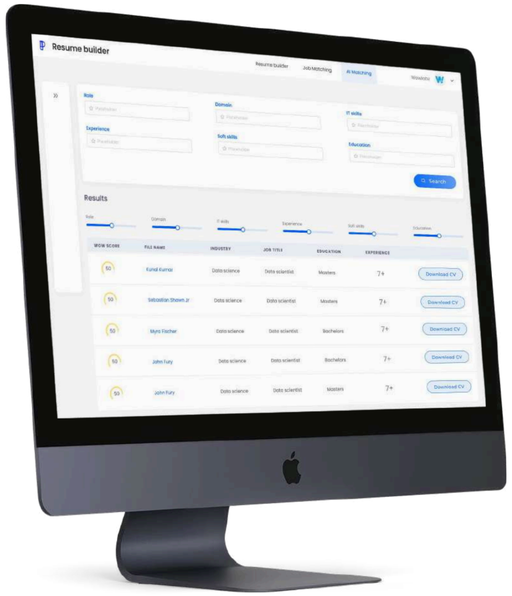
What Is NER ?

NER is the task of identifying and categorizing key information (entities) in text. This algorithm can extract different entities that are present in the resume like name, experience, education, skills, etc.

What Is Parser ?

Combining of object detection and NER models, we can transform an unstructured resume into a structured set of entities.

We employed object detection and NER algorithms to glean information from the resume. The object detection model identified the sections of the resume, and the NER models then extracted pertinent data from those sections.

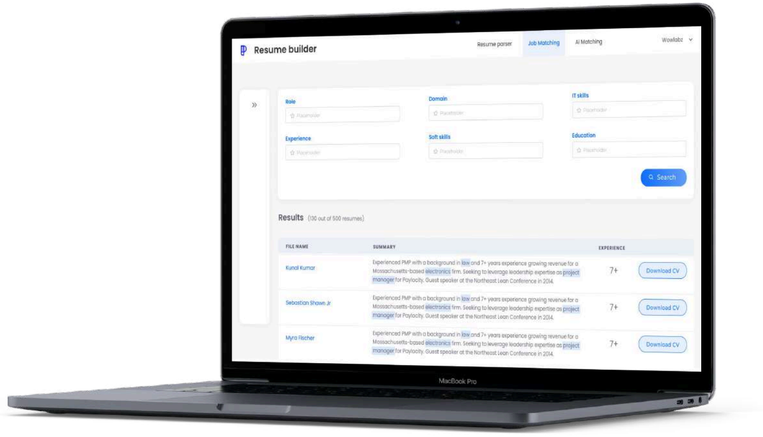


Section Parser

We used object detection methods to break down resumes into different sections like experience, education, personal details, etc. The detected cropped sections are converted to text using paddle OCR, an open source optical character recognition (OCR) toolkit.

Experience Parser

We utilized a NER model provided by spacy that utilizes a Convolutional Neural Network (CNN) to extract text features. This NER parser can identify entities such as company name, years of experience, location, job title and job description from an individual's experience.



Education Parser

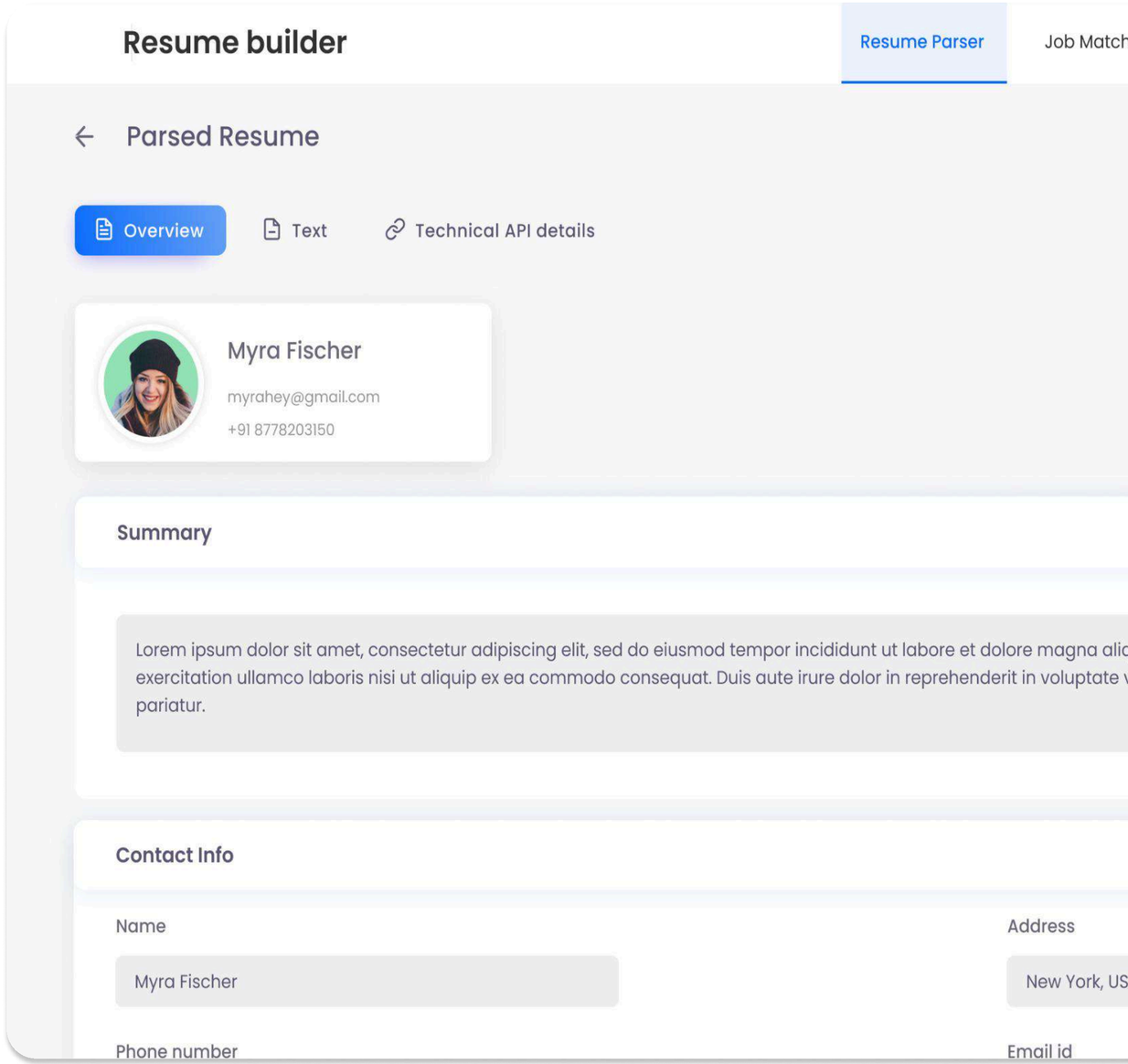
This NER parser is capable of recognising entities like college names, years, types of education, and scores from a person's educational background.

Skills Parser

The RapidFuzz library is used to compare the skills extracted from the section parser with the standard skills from the database, calculating the differences between the strings.

Personal details Parser

This NER parser is able to identify entities such as candidate address, date of birth, languages spoken and hobbies from resumes. Additionally, regular expressions were used to extract information like phone numbers, email addresses, and social media links.



Use cases

- Recruiting: Resume parsers can be used to quickly scan through large numbers of resumes and identify the most qualified candidates for a job.

- Automatically categorize resumes into different job categories, such as software engineering, marketing, and finance.

- Talent Acquisition: Resume parsers can be used to identify potential candidates for a job opening and quickly contact them.

- Automated Job Application: Resume parsers can be used to automatically fill out job applications with the information from a resume. This can save time and effort for job seekers.

Learnings

- The YOLOv5 model used in the resume parser can be improved by training on a large dataset of differently structured resumes.
- The NER model for experience, education and personal details parsing can be improved by training on a large dataset.

Thank You

You are welcome to our culture of shaping the future.

The WowLabz team is always working with our clients on solutions for new & challeging problems

wowlabz

Wow Internet Labz Pvt. Ltd.

1st Floor, Fremont Terraces, 3580, 4th Cross Rd,
HAL 2nd Stage, Doopanahalli, Indiranagar,
Bengaluru, Karnataka 560008

99163 19698

sales@wowlabz.com