SMART INTERNS IBM HACKATHON

AI Recruiter to Shortlist Candidate for Interview

BY

Andrea Lobo

Ignatius Almeida

Index

Introduction

Overview

Alongside the digital transformation over the past decade, developers have been working on Artificial Intelligence (AI) to help us with time consuming tasks, save us time and improve our daily lives. AI now helps us get to our destinations faster and predicts the weather better. These self-learning machines can analyze vast amounts of data in milliseconds and provide insights that make us smarter, more efficient, and better at the things we do every day. This is a key reason why AI is one of the hottest buzzwords in the business world today as executives search for ways to not only become more efficient, but also better at what they do. In talent acquisition, recruiting tools like applicant tracking systems have been able to automate mundane processes to help save time, they haven't necessarily been able to help HR teams work smarter. Simply put, these tools lack the ability to adapt on their own and provide the insights HR professionals desperately need. According to a recent survey by recruitment firm Hays, 92% of employers surveyed were seeing skills shortages that slowed their hiring and negatively affected their business. While AI can't magically give candidates skills to fill those gaps, it can help identify and automatically target more relevant candidates that are the closest fit. And that optimism is catching on; 80% of executives believe that AI recruiting can help make their hiring process more efficient.

Purpose

In a broad definition, AI enables computers to do things that, without it, would require human intervention like complex decision-making, problem solving and learning. This technology not only significantly improves the process and outcomes, it also takes into account things that are not planned for or known by humans, using data to make the best decisions when it comes to carrying out a task. Self-learning algorithms are required when the scope of data and the scale of the problem are just too big for human interaction.
One of the biggest challenges facing HR professionals today is finding the best talent to hire. This task has proven burdensome in the past, in part due to inefficient manual tasks that plague the recruiting process and the lack of access to the right data to make informed decisions. In fact, per research conducted by LinkedIn, 46% of recruiters and hiring managers have identified ''finding the right candidate'' as the biggest hurdle in hiring today.

Literature Survey

Existing Problem

The present study has been consider the following methodology adopted by (Jesson, Matheson, & Lacey, 2011) was implemented to serve the aim of the study. The process of methodology is based on six principles 1) Mapping the filed through a scoping review; 2) Comprehensive search; 3) Quality assessment; 4) Data extraction; 5) Synthesis, and finally 6) Write-up. The study was directed as follow. First, the research plan designed with having a interest in research questions and by involving of very key specific words and prepared a set of exclusion and inclusion standards. The present situation to establishing the current knowledge on artificial intelligence recruitment process and other related areas to have necessary to have a avenues for future research. The adopted following questions for the study 1) Which studies have been conducted that focused on AI recruitment process? 2) What were the main findings of these studies? and the study chosen to use various keywords to select very relevant studies, such as artificial intelligence, human resources, talent acquisition, recruitment trends, recruitment process and recruitment industry.This is constructed by using keywords and different numbers of hits were produced. In the third step abstract, papers contents were scanned manually and if it is relevant, to do that articles are read carefully and seen that it should be in the scope of the interest.

The study results covered both qualitative and quantitative insights of the artificial intelligence recruitment process, were presented.

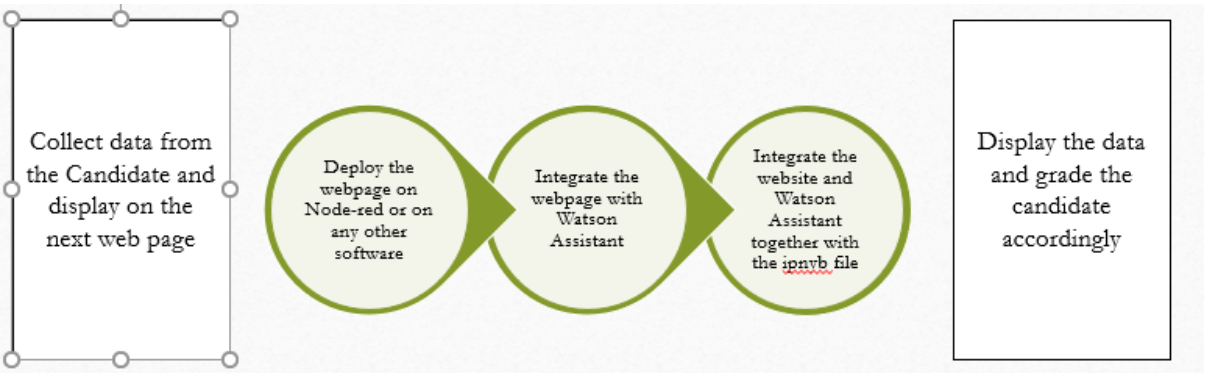
By having finalized number of the papers xxx through this designed the current knowledge of recruitment process and its related sub functions were established.

Moreover, a number other sub functions are spell out to address more to understand on AI recruitment process and it will be automatically highlighted. The present study embedded with associated for both theory & practice, when comes to the theoretical point of view, these paper given detailed discussion on the study of AI recruitment process, however it shows that, AI role in recruitment process, how it is cater other areas of recruitment process. In addition, the previous literature reviews brought together closer and considered artificial intelligence impact in recruitment process and its other area’s of such as screening candidate, building relationship, quality of hiring, scheduling, unbiased decisions, saving time, effort and money and so on.

Currently, a recruitment process is carried out mostly by human recruiters who personally sit and scan through CV’s, online profiles and other sources to find candidates. Recruiters conduct all initial contact, give feedback to rejected employees and conduct interviews with candidates (O’Donovan, 2019). As humans

have limited abilities, keeping up with all the tasks that is necessary is not an easy job, and usually requires lots of dedicated time from every individual recruiter. The problem that have been identified is that there are human limitations, such as biases, preconceptions and time restraints, which can hinder how effective a recruitment process ends up being (McRobert, Hill, Smale, Hay, & Van Der Windt, 2018). This is a problem as it, in turn, can lead an organization to lose the better fit candidates for a job as well as monetary value (Baron, Musthafa & Agustina, 2018). It has been identified that the methods of investigating technology-based recruitment are lacking and comes behind the current practice. Hence more in-depth empirical research must be conducted in the future with the regard of new technology allowing more flexibility and better access than before (Chapman & Webster, 2003; Searle, 2006). However, several years later the same problem is still here, since Marler and Fisher (2013) mention that the current literature is lacking the new technology-based recruitment methods that need to be fulfilled. In addition, the implications of new technologies for HRM are still somehow unclear for recruiters whether these new and efficient technologies entail challenges or opportunities to recruiters' work (Stone, Deadrick, Lukaszewski & Johnson, 2015; Bondarouk & Brewster, 2016). Since the current literature is still lacking the same problem as it was in the 2000s, a more in-depth understanding of the topic should to be conducted with the fact new technologies being a part of the recruiter's daily work.

Proposed Solution

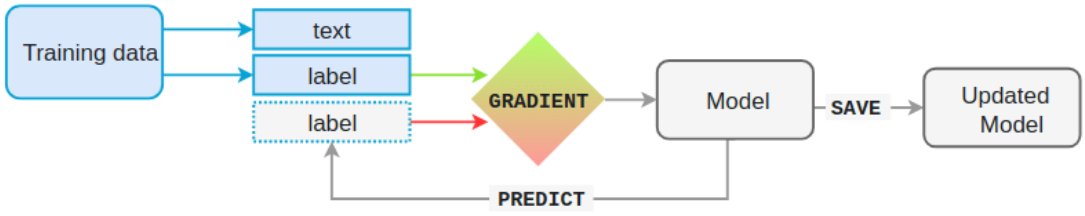


In our Project , we made a website where the candidate will fill in the details and will then have to caht with the Watson Assistant . We have also made a ipnyb file that would summarise the resume and display entities using NLP algorithms. For the summarisation of the resume we have used jupyter Notebook. We deployed the website on Heroku. We tried Node-red and Watson studio for the same.

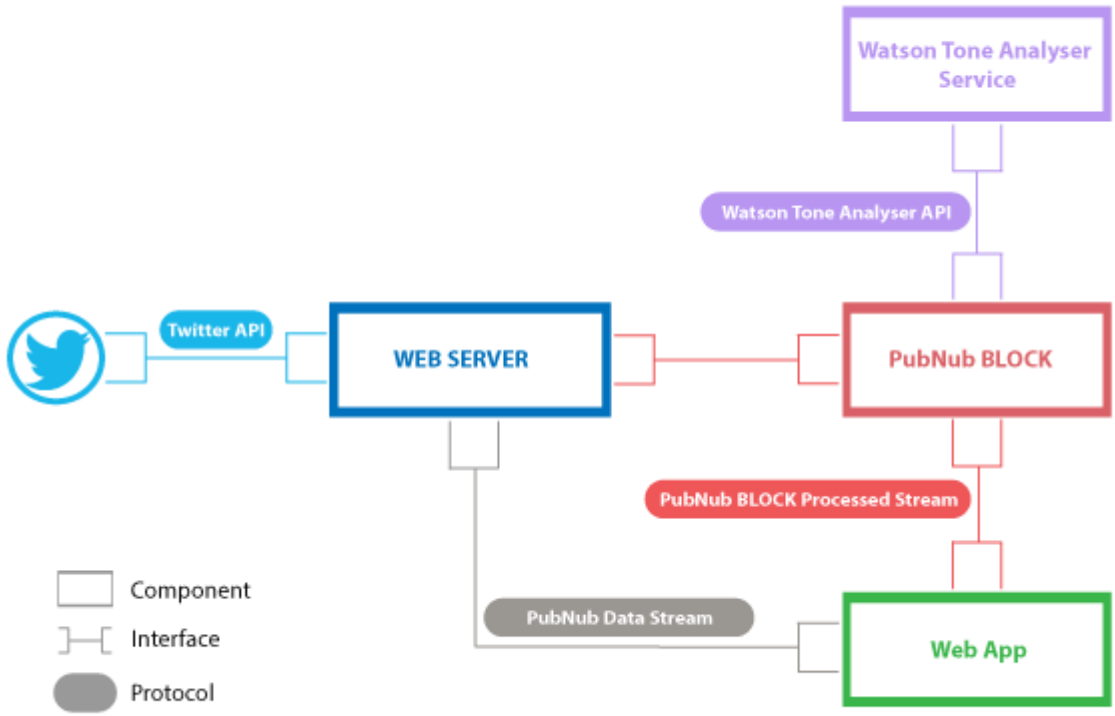
Theoritical Analysis

Block diagram

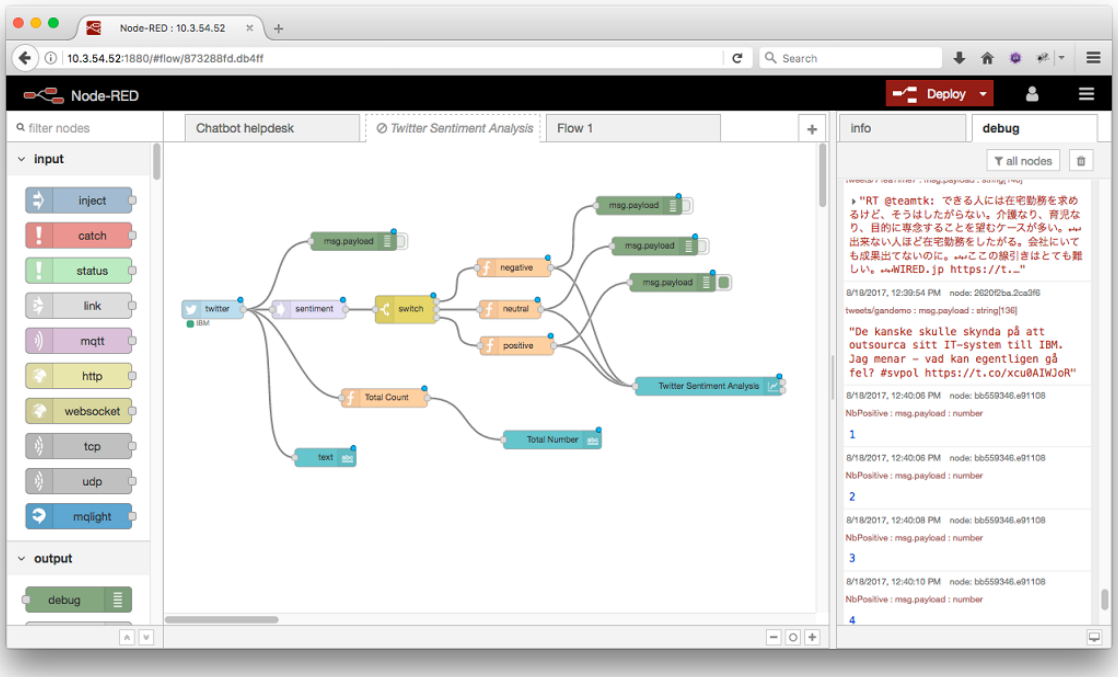
- Entity Recognition



- Using Cloud Services (Watson)



- Using Node-red



Hardware and software Designing

- Hardware
  - Speed:1.1 GHZ
  - RAM: 4 GB(min)
  - Hard Disk:40GB
- Software
  - Jupyter Notebook
  - Cloud IBM
  - Heroku
  - Sublime and Xampp

Experimental Investigations

Since , AI requires data, data, and oh, more data to work, a place where you're likely to see the most immediate benefit with adoption of AI-enabled recruiting solutions is in your key performance indicators (KPIs) like time to hire and cost per hire. These areas are probably where you are feeling the most pressure as well. Not only can AI help improve your day-day operations, which will have direct impact on your costs, it also offers new insights that will help you improve your overall strategy. You may even discover better KPIs to use along the way.

Cost per Hire

According to SHRM, the average cost-per-hire (CPH) is $4,129 per open position. This statistic keeps rising year over year as companies are forced to spend more on things such as posting on job boards and employment branding to fill their talent pipelines. In addition, HR departments are having to increase their staff to support all the extra work related to sourcing and screening to make the right hires in a timely manner. The only way to keep cost-per-hire in check without adversely affecting time-to-hire is to reduce costs or improve efficiency, and both are exactly what AI-enabled solutions aim to do. One example of how AI can help reduce time to hire where automation alone is not sufficient lies within the ATS. Applicant tracking systems have done a great job at automating straightforward tasks such as scheduling, providing status updates, triggering workflows, and more to speed up the hiring process.

Improving Quality of Hire

While time and cost are certainly major considerations, improving quality of hire (QoH) is still the king. More than 40% of recruiters and hiring managers cite it is the most important factor in hiring. To put a number to the QoH (also known as the QoH Index), you can attach a percentage to the following factors:

- Factor 1: New hires' productivity levels (which may include revenue or sales or meeting project goals within certain timeframes)
- Factor 2: Surveys both for new hires and hiring managers assessing performance (creating data using a scoring metric)
- Factor 3: Retention rates and turnover

Reducing Hiring Bias

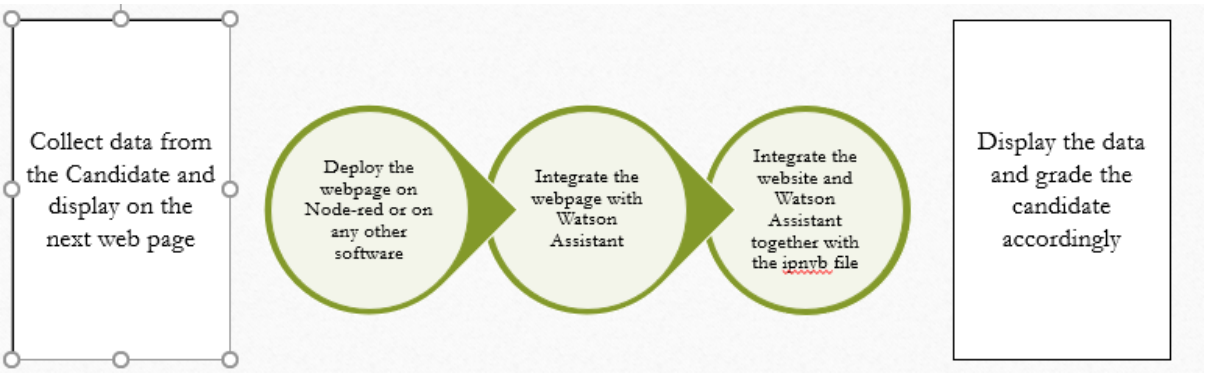
The data collected and reported by AI-enabled recruiting software doesn't just help study and improve the QoH metric; it also helps eliminate some of the most problematic human flaws in the hiring process. Eliminating hiring bias is one of the biggest challenges in recruiting today. Whether explicit or implicit, stereotypes and personal biases are something to which even the most conscientious recruiters can fall prey. AI helps level the playing field by allowing for blind applicant screening.

Reaching Qualified Passive Candidates

Passive candidates, or candidates who aren't actively seeking a new job, can actually be some of the best applicant pools. Reaching these candidates in the past meant a lot of resume database searching, cold-calling, outreach, and dead ends.

Although real-time matching technology can eliminate the hassles of keyword searching to identify passive candidates to target, some new solutions like Entelo are taking it one step further by aggregating a candidate's profile information and recent activity from different sources, such as public resume databases and social networks, to predict how 'receptive' the candidate may be on taking a new job as well.

Flow charts



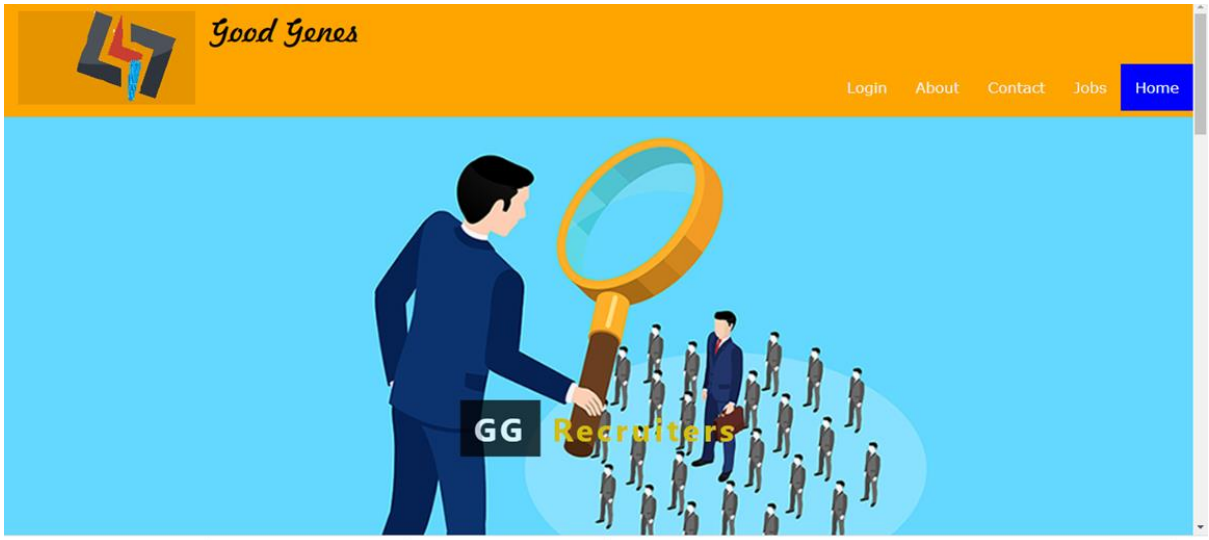
- To create a system for recruiting good candidates based on the resume and the personality test based on the chat with the Watson Assistant.
- For the Keyword Extraction from the resume , we have used Entity Recognition Algorithm of NLP(Natural Language Processing).

This project is created so as to select the best candidate and avoid human error

`

Result

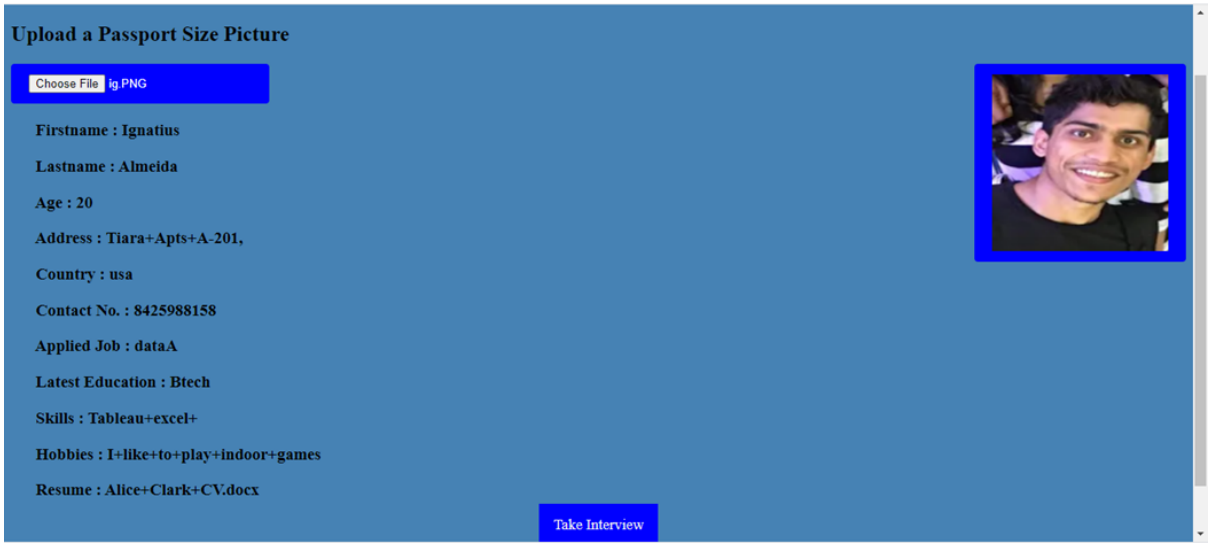
- Website Opening



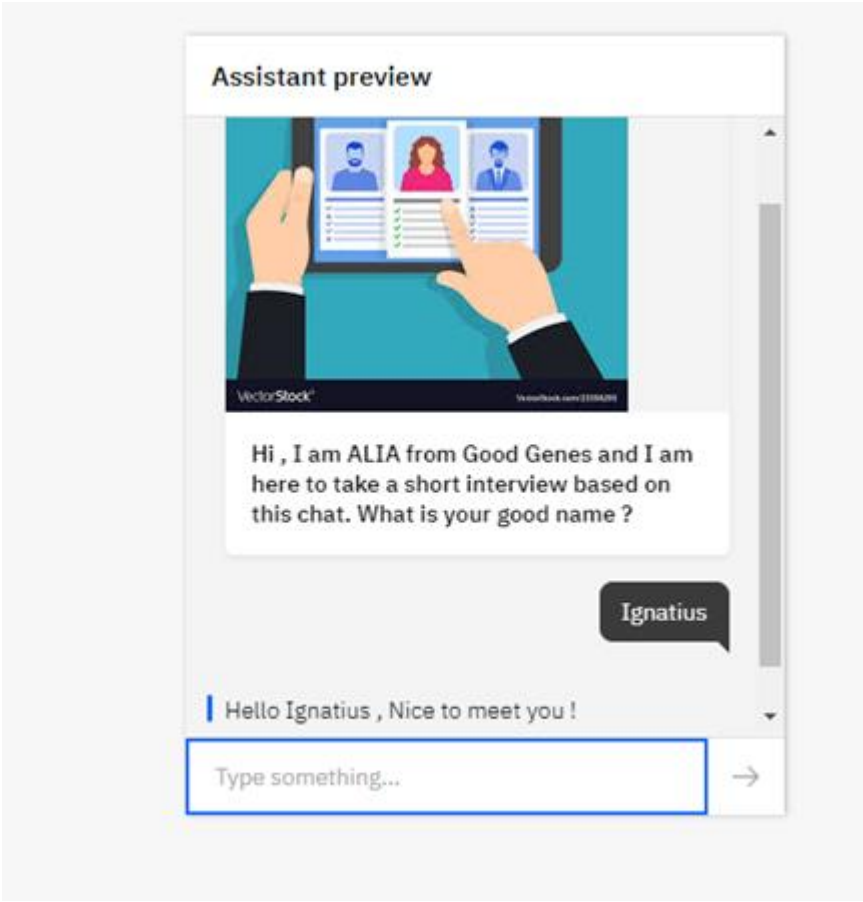
- Apply for Job Page

- The form Page

- Displaying Data



- Watson Assistant



- Keyword Extraction

jupyter resume Last Checkpoint: Yesterday at 4:35 PM (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run | Code

In [1]:
```
import spacy
import pickle
import random
```

In [2]:
```
train_data=pickle.load(open('train_data.pkl','rb'))
```

In [4]:
```
train_data[1]
```

Out[4]:
```
("Harini Komaravelli Test Analyst at Oracle, Hyderabad  Hyderabad, Telangana - Email me on Indeed: indeed.com/r/Harini-Komarav
e111/2659eee82e435d1b  > 6 Yrs. of IT Experience In Manual and Automation testing.  WORK EXPERIENCE  QA Analyst  Oracle - Test A
nalyst at Oracle, Hyderabad  Infosys Ltd -  Hyderabad, Telangana -  November 2011 to February 2016  Hyderabad from Nov 2011 to
Feb17 2016 > Worked in Tata Consultancy Services, Hyderabad from Feb 24 to Apr 11 2017 > Currently working as a Test Analyst a
t Oracle, Hyderabad  QA Analyst with 6 years of IT experience. Oracle  EDUCATION  MCA  Osmania University  B.Sc. in Computer Sc
ience  Osmania University  SKILLS  Functional Testing, Blue Prism, Qtp  ADDITIONAL INFORMATION  Area of Expertise:  > Familiar
with Agile Methodologies  > Having knowledge in Energy (Petroleum) & Health Care domains. > Involved in preparation of Test Sc
enarios. > Preparing Test Data for the test cases.  https://www.indeed.com/r/Harini-Komaravelli/2659eee82e435d1b?isid=rex-downl
oad&ikw=download-top&co=IN https://www.indeed.com/r/Harini-Komaravelli/2659eee82e435d1b?isid=rex-download&ikw=download-top&co=I
N  > Experienced in development and execution of Test cases effectively. > Experienced in Functional testing, GUI testing, Sm
oke testing, Regression testing and Integration Testing > Experienced in doing Accessibility testing of an application > Abili
ty to understand user Requirements, Functional and Design specifications, > Good knowledge of SDLC and STLC processes. > Decid
ing the Severity and Priority of bugs. > Experience in using Microsoft Test Manager & Oracle Test Manager as Test Management To
ols. > Having good experience in testing windows based & web based applications. > Involved in Client Interactions for review
s, issues and for any clarifications. > Web Services Testing > Writing Test Scripts in QTP, Testcomplete. > Creating Object R
epositories and Function Libraries in QTP. > Enhanced QTP scripts using VB Script. > Strong experience in working with Blue Pr
ism tool > Worked on different Environments like Windows Application & Web Application  Technical Skills:  ❑ Test Automation T
ools: Blue Prism, QTP 10.0, Testcomplete ❑ Test Management Tool: Microsoft Test Manager, Oracle Test Manager & JIRA ❑ Database
s: Oracle 10g, SQL Server  ❑ Operating Systems: Windows 7  Project 1: Title: Cadence Client: Baker Hughes  Technologies: Micro
soft Visual Studio and Microsoft Team Foundation Server  Client Background: An oilfield services company delivering focused eff
orts on shale gas and other oilfield services. It provides services, tools and software for drilling and formation evaluation,
well completion, production management, seismic data collection and interpretation.  Project Description: AUT (Application unde
r test) is the next generation revolutionary, robust, easy to use scalable well site data acquisition processing and interpreta
tion system for Client's Drilling Services to deliver services that meets cross divisional business requirements consistently.
Project 2:  Description: Paragon supports your entire care team with one tool that your clinicians need to help deliver the bes
t patient care. Designed by physicians, nurses, pharmacists and mid level providers that have a first-hand understanding of cli
nical workflow needs, Paragon clinical applications allow your caregivers to focus on what matters most; spending time caring f
or patients. Since Paragon is fully-integrated across all applications and built around a single patient database, information
entered anywhere in the system is immediately available to the entire care team. Immediate access not only helps clinicians mak
e better treatment decisions - it also helps promote patient safety. Paragon offers a broad suite of multidisciplinary clinical
software solutions together with anytime, anywhere access to the complete patient record.  Responsibilities:  • Performed Smoke
testing and Regression testing. • Involved in Generating and Executing Test Script using Quick Test Pro & Blue Prism • Usabilit
y and User Interface Testing. • Involved in Defect tracking and reporting the bugs using TFS • Participated in frequent walk-th
rough meetings with Internal Quality Assurance groups and with development groups. • Participated in client calls and clarifyin
g the doubts by having AT&T sessions • Involved in functional, regression and smoke testing to validate the application data ch
anges done in windows application • Certifying the build status by running the scripts as part of smoke testing  Project 3:  De
scription: Food & Beverages R&A: Easily manage business across multiple locations while reducing IT cost and complexity. Cloud-
based point-of-sale solutions enable centralized enterprise management with lower upfront costs and a smaller footprint.
Responsibilities:  • Performed Functional testing and Regression testing. • Involved in Generating and Executing Test Scripts u
sing Blue Prism tool and Open script • Involved in preparing bots using Blue Prism tool. • Accessibility testing of the web app
lication • Involved in defect tracking and reporting the bugs using JIRA • Webservices testing by calling API's to export the d
ata',
 {'entities': [(2275, 2281, 'Companies worked at'),
   (2235, 2241, 'Companies worked at'),
   (1603, 1609, 'Companies worked at'),
   (667, 703, 'Skills'),
   (638, 658, 'College Name'),
   (612, 637, 'Degree'),
   (591, 611, 'College Name'),
   (587, 590, 'Degree'),
   (568, 574, 'Companies worked at'),
   (526, 536, 'Designation'),
   (515, 524, 'Location'),
   (507, 513, 'Companies worked at'),
   (491, 503, 'Designation'),
   (429, 438, 'Location'),
   (352, 361, 'Location'),
   (296, 305, 'Location'),
   (270, 279, 'Location'),
   (262, 268, 'Companies worked at'),
   (246, 258, 'Designation'),
   (238, 244, 'Companies worked at'),
   (226, 236, 'Designation'),
   (177, 207, 'Designation'),
   (150, 155, 'Years of Experience'),
   (54, 63, 'Location'),
   (43, 52, 'Location'),
   (35, 41, 'Companies worked at'),
   (19, 31, 'Designation'),
   (0, 18, 'Name')]}
```

In [5]:
```
nlp= spacy.blank('en')
def train_model(train_data):
    if 'ner' not in nlp.pipe_names:
        ner=nlp.create_pipe('ner')
        nlp.add_pipe(ner, last=True)
    for _, annotation in train_data:
        for ent in annotation['entities']:
            ner.add_label(ent[2])

    other_pipes=[pipe for pipe in nlp.pipe_names if pipe!='ner']
    with nlp.disable_pipes(*other_pipes):
        optimizer=nlp.begin_training()
        for itn in range(10):
            print("Starting Iteration :" +str(itn))
            random.shuffle(train_data)
            losses={}
            index=0
            for text, annotations in train_data:
                try:
                    nlp.update(
                        [text],
                        [annotations],
                        drop=0.2,
                        sgd=optimizer,
                        losses=losses)
                except Exception as e:
                    pass

            print(losses)
```

In [6]:
```
train_model(train_data)
```

```
Starting Iteration :0
{'ner': 15339.812839700744}
Starting Iteration :1
{'ner': 14496.669588501514}
Starting Iteration :2
{'ner': 11593.916761988423}
Starting Iteration :3
{'ner': 10195.567154476756}
Starting Iteration :4
{'ner': 9055.190297840223}
Starting Iteration :5
{'ner': 10020.400477472824}
Starting Iteration :6
{'ner': 8480.408631914324}
Starting Iteration :7
{'ner': 9073.597988862422}
Starting Iteration :8
{'ner': 7127.256508364303S}
Starting Iteration :9
{'ner': 6221.750934825815}
```

In [7]:
```
nlp.to_disk('nlp_model')
```

In [8]:
```
nlp_model=spacy.load('nlp_model')
```

In [ ]:
```
```

In [19]:
```
train_data[1][1]
```

Out[19]:
```
{'entities': [(2301, 2380, 'Skills'),
  (2088, 2131, 'College Name'),
  (2080, 2086, 'Degree'),
  (2042, 2067, 'Companies worked at'),
  (2025, 2040, 'Designation'),
  (1445, 1470, 'Companies worked at'),
  (1094, 1134, 'Email Address'),
  (809, 834, 'Companies worked at'),
  (792, 807, 'Designation'),
  (616, 680, 'Skills'),
  (127, 142, 'Designation'),
  (83, 123, 'Email Address'),
  (56, 61, 'Location'),
  (29, 54, 'Companies worked at'),
  (11, 26, 'Designation'),
  (0, 10, 'Name')]}
```

In [14]:
```
doc=nlp_model(train_data[1][0])
for ent in doc.ents:
    print(f'{ent.label_.upper():30}-{ent.text}')
```

```
NAME                           -Arpit Jain
DESIGNATION                    -Quality Analyst
COMPANIES WORKED AT            -ThoughtWorks Technologies
LOCATION                       -Pune
EMAIL ADDRESS                  -indeed.com/r/Arpit-Jain/3714fe32f08b03a9
DESIGNATION                    -Quality Analyst
DESIGNATION                    -Quality Analyst
COMPANIES WORKED AT            -ThoughtWorks Technologies
COMPANIES WORKED AT            -Infosys Ltd
COMPANIES WORKED AT            -ThoughtWorks Technologies
DESIGNATION                    -Quality Analyst
COMPANIES WORKED AT            -ThoughtWorks Technologies
DEGREE                         -B.Tech
COLLEGE NAME                   -Jaypee Institute Of Information Technology
```

In [20]:
```
import sys , fitz
fname='Andrealobo.pdf'
doc=fitz.open(fname)
text=""
for page in doc:
    text=text+ str(page.getText())
tx=" ".join(text.split('\n'))
print(tx)
```

```
Andrea Lobo andrealobo1399@gmail.com +91 8425088158 Mumbai EDUCATION Bachelor of Technology (B.Tech), Information and Communica
tion  Technology (2017 - 2021) st. francis institute of technology XII (Senior Secondary), Science Year of Completion: 2017 HSC
Board (Shri Tp Bhatia) X (Secondary) Year of Completion: 2015 SSC Board (Carmel Of St. Joseph) INTERNSHIPS Web Software Kubis S
quare (Virtual) May 2020 - Present Web Development Dight International Chambers (Virtual) May 2020 - Present Data Science Cloud
Counsellage (Virtual) Mar 2020 - Present I have worked as a data scientist on live projects with learning paths. TRAININGS Web D
evelopment Anudip Foundations (Online) Apr 2020 - Present PROJECTS sportiqo Jul 2019 - Nov 2019 Sportiqo is a college Internet
programming project which is a website based on sports collection Python project Jan 2019 - May 2019 Hospital management syste
m in which every individual is given importance . This project was made as my Python project in the 4th semester SKILLS PHP Be
ginner Python Beginner HTML Beginner Java Beginner JavaScript Beginner JavaFX Beginner C Programming Intermediate Adobe Photosh
op Beginner
```

In [21]:
```
doc=nlp_model(tx)
for ent in doc.ents:
    print(f'{ent.label_.upper():30}-{ent.text}')
```

```
NAME                           -Andrea Lobo
DESIGNATION                    -Java Beginner JavaScript Beginner JavaFX Beginner C Programming
```

In [ ]:
```
```

Codes:

```
#!/usr/bin/env python
# coding: utf-8

# In[1]:


import spacy
import pickle
import random


# In[2]:


train_data=pickle.load(open('train_data.pkl','rb'))


# In[4]:


train_data[1]


# In[5]:


nlp= spacy.blank('en')
def train_model(train_data):
    if 'ner' not in nlp.pipe_names:
        ner=nlp.create_pipe('ner')
        nlp.add_pipe(ner, last=True)
    for _, annotation in train_data:
        for ent in annotation['entities']:
            ner.add_label(ent[2])

    other_pipes=[pipe for pipe in nlp.pipe_names if pipe!='ner']
    with nlp.disable_pipes(*other_pipes):
        optimizer=nlp.begin_training()
```

```
        for itn in range(10):
            print("Starting Iteration :" +str(itn))
            random.shuffle(train_data)
            losses={}
            index=0
            for text, annotations in train_data:
                try:
                    nlp.update(
                    [text],
                    [annotations],
                    drop=0.2,
                    sgd=optimizer,
                    losses=losses)
                except Exception as e:
                    pass

            print(losses)

# In[6]:
train_model(train_data)
# In[7]:
nlp.to_disk('nlp_model')
# In[8]:
nlp_model=spacy.load('nlp_model')
 # In[ ]:




# In[19]:


train_data[1][1]


# In[14]:


doc=nlp_model(train_data[1][0])
for ent in doc.ents:
```

```
    print(f'{ent.label_.upper():{30}}-{ent.text}')


# In[20]:


import sys , fitz
fname='AndreaLobo.pdf'
doc=fitz.open(fname)
text=""
for page in doc:
    text=text+ str(page.getText())
tx=" ".join(text.split('\n'))
print(tx)


# In[21]:


doc=nlp_model(tx)
for ent in doc.ents:
    print(f'{ent.label_.upper():{30}}-{ent.text}')
```

Advantages & Disadvantages

❖ Advantages

- Improved Quality of Candidates.
- Automate Tedious Manual Tasks.
- Better Experience for Candidates.
- An optimized **Recruitment** Process.
- Cost Effective **Hiring**.
- Reduced Time to Hire.
- No more 'talent waste'

❖ Disadvantages

- Applicants may not love interacting more with things like chatbots (recruitment bots) and the like. The lack of personal interaction may turn some people away.

- Putting hard rules in place to screen applications may create too much rigidity, thus possibly overlooking candidates who have nontraditional backgrounds but would otherwise be perfectly qualified and able to do the job.

- AI doesn't always eliminate biases because it still has to be taught what to look for—so biases can still come into the process inadvertently.

Applications

AI for recruiting has several potential applications for automating high-volume, repetitive tasks such as resume screening and pre-qualifying candidates.

1. Intelligent screening software

Intelligent screening software automates resume screening by using AI (i.e., machine learning) on your existing resume database.

2. Recruiter chatbots

Recruiter chatbots are currently being tested to provide real-time interaction to candidates by asking questions based on the job requirements and providing feedback, updates, and next-step suggestions.

3. Digitized interviews

Online interview software has been available for a while, but today's technology claims to use AI to assess candidates' word choices, speech patterns, and facial expressions to assess his or her fit for the role and possibly even the organization and its culture.

Conclusion

AI will change the recruiter role through augmented intelligence which will allow recruiters to become more proactive in their hiring, help determine a candidate's culture fit, and improve their relationships with hiring managers by using data to measure KPIs such as quality of hire.In this project, we have added Cloud IBM resources usefully and have learned and gained knowledgeable insights on its services

Future Scope

- We can make use of the Watson Studio to gather images of the Candidate and then Classify them.
- The Chat bot can be made as available as a video and can have a voice assistant.
- We could have integrated the python ipnyb with Flask Python Libraries for effective Interactivity

Biblography

1. https://ideal.com/ai-recruiting/

2. https://hrdailyadvisor.blr.com/2019/06/05/cons-of-using-ai-in-the-recruiting-process/#:~:text=Using%20AI%20in%20the%20Recruiting%20Process%3A%20Cons,may%20turn%20some%20people%20away.

3. https://www.pandologic.com/recruiting-with-ai/

4. https://chatbotsmagazine.com/how-to-develop-a-chatbot-from-scratch-62bed1adab8c