Interviewbot

Machine interviewing for more efficient applicant screening

Prepared by:
Anupam Misra
SHAHRUKH GOUHAR
Radha Rathore
Mansoor Ali Shaik

1. Business landscape

On average, for every open position, at least ten candidates are applying. Automated resume screening applications already exist. We wanted to automate the next stage of the hiring process: interviews. We read about chatbots and conversational AI to understand how to conduct interviews. Since we are studying and interviewing with data science companies, we decided it would be interesting to make an interview bot to interview candidates on data science questions.

2. Business challenge

There is a large gap between the supply and demand in the data science sector in the country. Many companies use automated resume screening. Even then the hiring managers are faced with the problem of inviting a large no. of prospective applicants to interviews.

Some companies do telephonic interviews as the first line of screening. More often than not, human emotions take precedence over logic which makes the whole process biased. We have designed an interview system aimed to replace this stage of screening for a more fair evaluation.

3. Data used

3.1 Gathering and structuring the data

Data is the main requirement for any data science project. As we were going to conduct interviews, we built a repository of question-answer pairs. We created a database of questions from Statistics, Linear regression, Logistic Regression, KNN, SVM, KMeans, Decision Tree, and Naive Bayes. We kept questions of three difficulty levels from the aforementioned topics. In total, we created around 250 questions. The question-answer pairs were either web-scrapped or entered manually. The answers

were scrutinized to remove non-contextual words and include specific relevant keywords.

3.2 The need for multiple answers

During several rounds of testing, we realized that some questions could be answered with multiple keywords. Hence, up to three answers were written for those questions wherever applicable. During scoring, the user's answer closest to either of the three answers was considered for scoring.

3.3 Creating a separate test dataset

We tried different approaches like sentence-transformers, LSA, and neural networks for scoring the user answers. As we are working with a very niche text vocabulary we had to separately create a test dataset comprising question-answer pairs from the original dataset. However, in this dataset, we manually added a new attribute called user-answer to indicate user answers of different levels of correctness against each question. Expected scores were also allotted against each user-answer variation to train and evaluate the different scoring models.

4. Describe the solution

Our application is designed to help businesses redefine their interview process by enhancing and optimizing it. The aim is to reduce time, cost, and effort by making the interview screening process smarter, faster, and fair.

The first stage of any interview process begins with preparing a proper interview schedule for numerous candidates, which is a long, tedious, and error-prone process. This is solved smartly by our application by parallelizing the process. As the candidates are evaluated by a machine, there is no scope of bias. This will result in better screening of candidates. The evaluation process is based on advanced machine learning logic which compares users' answers with standard answers and marks them based on the level of similarity. The scoring is done by a Random forest model being fed similarity scores from BERT and a GRU network with custom-trained word embeddings.

Detailed reports after the interview would be helpful in better analysis of candidates' performance and making the hiring decision.

5. Impact analysis

According to SHRM, the average time it takes to hire a new employee is 36 days. The impact of our application can be measured through the following metrics:

1. Speed to hire ratio - The speed of hiring will increase.
2. Average time to fill a position - The average time to fill an open position will go down
3. Interviewee ratio - The no. of interviews taken per position
4. Recruitment cost per hire - The recruitment cost per hire will go down