Scientific Progress Report

of BSc students of 4 year

6B (06101) Information Systems

Use of Big Data to improve HR practices

Team members:

- Adilkhan Kuchkarov - 190103427
- Baltabay Sarsenbay - 190103251
- Madiyar Nigarbay - 190103327
- Rakhat Alpysbai - 190103345
- Tamerlan Bulatkaliyev – 190103483

Supervisor: Ualikhan Sadyk

Content

Keywords: big data, HR practices, data analysis, machine learning, text processing, people analytics, neural network

Abstract

The rest of the report is structured as follows.

- Firstly, we will look at the explanation of the neural network.
- Next, we will look at the business logic of our project.
- After that, we will look at specific tools to achieve our goal.

The main requirements for the development of this project were a good knowledge of the python programming language and its library for working with big data, machine learning, and neural networks, and basic knowledge of text processing (NLP)

Introduction

In all possible new technologies, people are being replaced. And including companies resort to unusual individual selection for employees of the personnel department. Machine learning has conquered many fields of activity in a short time and has developed in a secure way. In our report, we will tell you how you can improve the practice of HR employees.

Aim and objectives of the research

The purpose of this report is to offer a shared vision of how we can improve HR practices for the personal selection process using neural networks trained on big data for prediction.

Tasks to solve this:

1. Collect a large dataset consisting of resumes or data of applicants. You can parse data from job search sites, or look for ready-made datasets from data analysis sites (e.g. kaggle.com).
2. Learn how to create a machine learning model that, based on these data, will display a verdict with close to 100% accuracy.
3. Build a business plan for our service.
4. Create a website for the provision of services for the selection of employees.

Business logic

In short, the business logic of our project is to make a service for predicting the best candidate listed for the position presented by the company.

We will create a neural network that we will train to find the best candidate based on the data provided.

The company that contacted us provides a list of these candidates and the desired position.

What kind of data is valuable for the forecast?

We discussed this issue and narrowed it down to these data, which in our opinion are more valuable: career level (initial, intermediate, intermediate, senior, or executive level), discipline, region, the coincidence of job roles

Tools to achieve our goal

Machine learning can be implemented using programming languages such as Matlab, Java, and Python. Our choice was between Python and Java, but in the end, we decided to use Python because it has a huge number of libraries to solve our problems. There are libraries such as:

Pandas to work with big datasets, *NumPy for math*, *scikit-learn*, *Keras for machine learning*, and *Django for website creation.*

Literature Review

It's no secret that wrong decisions when choosing employees will lead a business to failure. Therefore, the role of recruiters is very important. But the human factor can lead to wrong decisions. The reasons may be different. All people do mistakes. To avoid the human factor in recruiting, we can use the power of machine learning by analyzing data.

The problem with using machine learning is the small amount of worthwhile, structured data for analysis and a large variation of the profession, where the key skills differ.

Many studies have used a small set of data or functions. This can lead the model to serious errors or not very accurate answers when the verdict is given.

It is important to select several key functions that are more important when choosing employees. Excess can also have a bad effect on the final result. In some studies, leadership, education, work experience, or psychological assessments, age, and marital status of the candidate are chosen as important functions.

Understanding why the model showed a verdict and by what criteria it showed it is important to use a model that gives interpreted information. The non-interpreted model represents exclusively ranked estimates or forecasts.

In conclusion, we realized that the solution to this problem consists in choosing key functions, and creating a machine learning model that gives interpreted information and a large amount of useful structured data, which is also one big task to take.

References

1. HR Analytics Neural Network in Keras: https://www.kaggle.com/code/neilslab/hr-analytics-neural-network-in-keras/notebook
2. Artificial Neural Network system applied to Human Resource Management: https://upcommons.upc.edu/handle/2117/109555
3. Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming: https://www.sciencedirect.com/science/article/pii/S0167923620300452
4. Human resource management and performance: A neural network analysis: https://www.sciencedirect.com/science/article/abs/pii/S037722170600405X
5. Application of Machine Learning Algorithms to an online Recruitment System:
6. https://d1wqtxts1xzle7.cloudfront.net/76620138/download-with-cover-page-v2.pdf?Expires=1667041953&Signature=S1oXzi8Iagc4d1KmBXYuPwwav4VKDPL9je9wy~IWVUo-2XAUrud83a~Ta3QfRmcd-SKmrmTLS7cZ15sQxgaDqiupo6N0uJKE-knh2KLyw13H0qmKOxbE~Mm05ZAPcUzl3RnRo0E4vT7cNZftcTts1wYCQ-P6KIlwGUoAicQ13lxgTywqzNldegqATZhaJMTicuWjQ7jpbmo-Qt4eg25JvYgQcAyAfC0yDLhxJnU97LZm7-KxT0DysWB2Unok7xg~zprNvK8bkU70i00LAsEJdcAPi6nQ-Y2bpL3DguyQuLzSRFbvJVb7OFgw0yBIdYAzI3a4wtWoeCaYPTpIObqoeQ__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA
7. Recruitment Prediction using Machine Learning: https://ieeexplore.ieee.org/abstract/document/9276955
8. Recruitment in the times of machine learning: https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-6e4acda3-2c53-4715-9dee-c9b3231cf25d
9. Ethical Considerations in AI-Based Recruitment: https://ieeexplore.ieee.org/abstract/document/8937920
10. Student career prediction using advanced machine learning techniques: http://tayberrytech.com/download/Agriculture.pdf