INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078

(Volume 8, Issue 3 - V8I3-1357)
Available online at: https://www.ijariit.com

Employee's Salary Prediction

Tiasa Mukherjee
tmukherj@gitam.in
Gitam School of Science, Gitam University, Visakhapatnam, Andhra Pradesh

MS. B. Satyasaivani
svedula@gitam.edu
Gitam School of Science, Gitam University, Visakhapatnam, Andhra Pradesh

ABSTRACT

In machine learning is a technology which is a software program to predict or detect more accurate developing the computer program that can use data and use it to learn itself. In today's life prediction is very trendy technology which is predicts very accurately just like a human, to solving most of the prediction and detection problem. In this paper we are proposing employee salary prediction with some key attribute and machine learning algorithm.

Keywords: *Machine Learning, Supervised Learning, Linear Regression Algorithm*

1. INTRODUCTION

In today's employee are changed their company very frequently is the salary of the company. Employee changed their company one of the causes is not getting desired salary, sometimes company also loss their own for giving desired salary of the employee. Actual today's world is full of competition world and all of them has higher expectations and goal in their own. But can't able to give everyone's expected salary because we should be a system to calculate the capability of their employee to expected salary.

In this paper the main aim is predicting the salary of the employee based on their year of experience and hard works. In this system we are mainly working on the linear regression algorithm in supervised learning in machine learning.

In supervised learning the name denoted that a supervisor or a teacher. Mainly supervised learning algorithm are working on the labelled data means that some input data are already give the correct output. Supervised learning algorithm is the procedure which is providing input data as well as the correct output data.

Linear regression is an algorithm in machine learning which is based on supervised learning. Linear regression algorithm is mainly defining the linear relationship between a dependent variable and the one or more independent variable is basically a linear regression. The linear regression model represents the sloped straight line representing relationship between the variables.

1.1 Methodology

In this paper we are mainly predicted the employee's salary based on their year of experience. The methodology having the different phases like: data collection, data cleaning, feature engineering, data visualization, splitting data into training data and testing data, train data, test data, training data visualization, test data visualization, accuracy, output.

Step1: Firstly, we have to collect the dataset in the various open-source platform.

Step2: Secondly, we have to perform the data preprocessing like EDA (Exploratory Data Analysis) in the given dataset. In the EDA step we have to check that any null values are there or not, then check the information of the data, then describe the which shows the mean value, standard deviation value, maximum value etc.

Step 3: Thirdly, after completing the EDA part we have to visualize the data in the two-attribute data YearofExperience and the salary using the Scatterplot diagram.

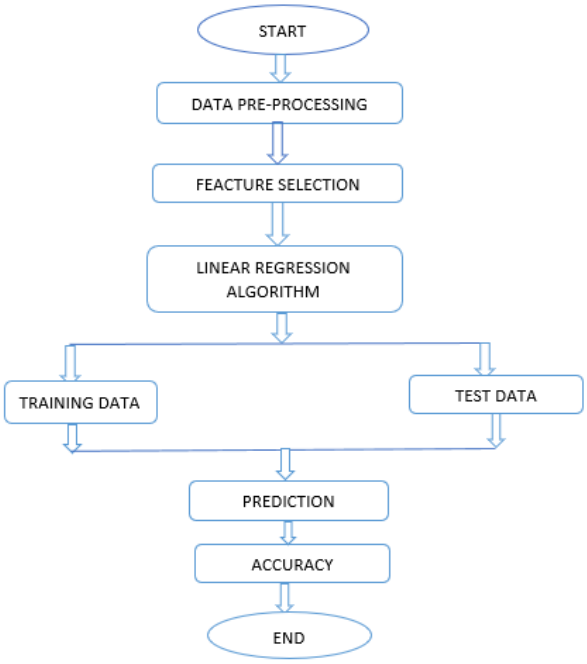

FIG: Architecture Diagram

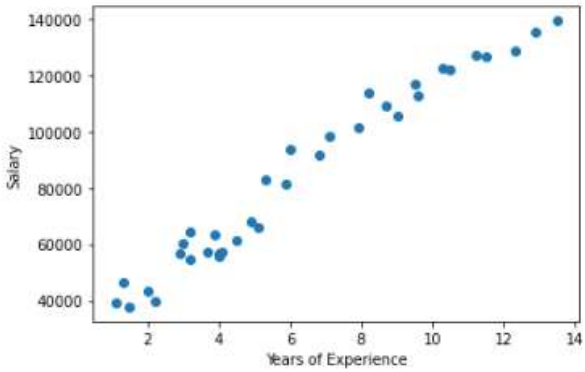

Fig: data visualization

Step 4: Fourthly, we have to preparing the data, divide the data into the dependent and independent variable. Then split the data into training data and test data.

Step 5: After splitting the data perform the linearRegression model with by default parameters and trained LinearRegression model with training data. And test the data then visualize the predicted data and actual data. Then calculate the difference between the actual salary value and predicted value.

Step 6: After that we have to visualize the training data, and draw the best fit line and plot all the training points of the training data and see the bias.

Fig: Training data visualization

Step 7: Visualize the test data, draw the best fit line and plotting the testing points of the test data and set bias.

Fig: Test data visualization

Step 8: In this step we have to calculate the accuracy of the model. This model accuracy is 97%. After that we have to check that my prediction is properly working or not.

2. LITERATURE REVIEW

In this paper [1], model is build using a regression model with the dependent variable means the salary and the other independent variable and in the same problem can be modeled in classification from the category of a given dataset.

In this paper [2], A salary prediction method based on data mining techniques is discussed in this paper. Using the Electronic copy available at: https://ssrn.com/abstract=3866758 graduated student results, a 10-fold cross-validation experiment was carried out. The result shows that, compared to KNN and Naive Bayes, the decision tree (J48) provides the best result.

In this paper [3], proposed prediction model using Decision tree technique with seven features. Moreover, the result of the system is not only a predicted salary, but also the 3-highest salary of the graduated students which share common attributes to the users. To test the system's efficiency, they set up an experiment by using 13,541 records of actual graduated student data. The total result in accuracy is 41.39%.

In this paper [4]," Salary Predictor System for Thailand Labour Workforce using Deep Learning" - used Deep learning techniques to construct a model which predicts the monthly salary of job seekers in Thailand solving a regression problem which is a numerical outcome is effective. We used five-month personal profile data from well-known job search website for the analysis. As a result, Deep learning model has strong performance whether accuracy or process time by RMSE 0.774 x 104 and only 17 seconds for runtime.

In this paper [5], they evaluate the importance of the features that can be used to forecast wages, after examining key aspects of the job market. Results suggest that these attributes contribute substantially to the final compensation perceived by workers, such as experience, work security, or certain job positions.

In this paper [6], Author developed such that there are initial feature vectors combined with log-based feature reduction. 1. A negated meaning is indicated by this semantic feature. 2. Quantification that is universal. 3. trigram of a part of speech 4. A noun phrase that starts with a pronoun and ends with a punctuation character.

3.CONCLUSION

Employee salary prediction using machine learning process helps us to predict the employee's salary based on there year of experiences and there are many algorithms to predict this. Some of the algorithms are SVM, Logistic Regression and KNN. But even though we have many algorithms correct prediction is necessary to detect the employee's salary so that necessary precautions can be taken by the individual and incorrect prediction might leads to consequences for the individual.

So, in order to predict correctly the particular algorithm needs to be accurate and hence we can have a better performance on predicting the result. Linear regression is one of the algorithms with good accuracy of 97% and can be a good algorithm that would be able to predict employee's salary more effectively than other algorithms. This can be better shown by importing the dataset and calculating its accuracy score and even plots are drawn for better visualization of the data.

4.ACKNOWLEDGEMENT

First and foremost, I would like to thank the Almighty who always showers his blessing and gives me opportunity to do all good work, without his will it would not have been possible. I sincerely thank the Department of Computer Science for providing me good platform and necessary facilities.

I express my gratitude to Prof. M. Saratchandra Babu, our respected principal for the facilities provided by him throughout the course.

I would like to thank our respected Head of The Department Dr. T Uma Devi (Associate Professor) for giving me such a wonderful opportunity to expand my knowledge for my own branch.

I would like to express my gratitude to my supervisor, Associate Professor, Ms.B. Satyasaivani, Department of Computer Science, GITAM (Deemed to be University), Visakhapatnam, under whose supervision I was able to complete my project. I am really thankful to her for guiding and correcting various documents of mine with attention and care.
I thank one and all who have contributed directly or indirectly to my project.

5.REFERENCES

[1] Mangui Wu, Shunmin Shu," Top Management Salary, Stock Ratio and Firm Performance: A Comparative Study of State-owned and Private Listed Companies in China
[2] Navyashree M, Navyashree M K, Neetu M, Pooja G R, Arun Biradar "Salary Prediction in It Job Market", International Journal of Computer Sciences and Engineering Vol.-7, May 2019.
[3] Pornthep Khongchai, Pokpong Songmuang, "Improving Students' Motivation to Study using Salary Prediction System" 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)
[4] Phuwadol Viroonluecha, Thongchai Kaewkiriya," Salary Predictor System for Thailand Labour Workforce using Deep Learning" The 18th International Symposium on Communications and Information Technologies (ISCIT 2018)

BIBLIOGRAPHY

Tiasa Mukherjee
Gitam School of Science, Gitam University, Visakhapatnam, Andhra Pradesh, India

MS. B. Satyasaivani
Gitam School of Science, Gitam University, Visakhapatnam, Andhra Pradesh, India

