Proposal for Showing the Value of Machine Learning Model for Salary Prediction:

Problem and Market Size:

Our machine learning application solves **the problem** of accurately predicting salaries based on various factors such as education level, work experience, job title, industry, location, etc. It can help both employers and employees make informed decisions about salary negotiations and job offers. The potential and prospects for this product are remarkable, as accurate salary predictions are crucial for companies to attract and retain talent, and for individuals to make informed career decisions.

The **market size** for this product is significant, as it can be used by businesses and organizations of all sizes in various industries. According to our research, the global market for talent management software in 2023 is projected to grow from USD 7.93 billion in 2022 to USD 9.82 billion by 2028, at a CAGR of 12.1%12. The market growth is driven by the increasing demand for talent management solutions across various industries, the rising adoption of cloud-based and web-based applications, and the growing popularity of social media platforms. The market is segmented by solution, deployment mode, organization size, industry vertical, and region. North America is expected to hold a substantial market share, while Asia Pacific is expected to witness a high growth rate12. Some of the key players in the market are IBM Corp., Oracle Corp., SAP SE, Peoplefluent Inc., CornerStone OnDemand Inc., and others. (reference: [Talent Management Software Market Size, Growth, Analysis 2023-2028 (expertmarketresearch.com)](expertmarketresearch.com))

Results and comparison with other models using Performance Metrics:

We trained and evaluated our machine learning model for salary prediction using various metrics and techniques. Our model - (Random Forest) was trained on our dataset and achieved a high accuracy of 89.01% using metrics such as R-squared (Coefficient of determination) and RSME (root mean squared error). This outperformed the regression models such as SVC (Support Vector Classifier) and Decision Tree Classifier, respectively. Thus, our machine learning approach offers a robust and reliable solution for accurate salary prediction.

Monetary Value and Risks:

The monetary value of our application is significant, as it can help businesses save costs by providing accurate salary predictions. For example, if a company has 100 employees and our model can save 5% of their salary costs, this will result in savings of approximately $50,000 per year. If our model is used to negotiate salaries, it can also result in better retention of employees, which can save costs associated with recruiting and training new employees.

The risks associated with our application are minimal, as the accuracy of our predictions has been validated through rigorous testing and comparison with other models. However, there is a risk that the predictions may not always be accurate due to factors such as incomplete or inaccurate data inputs, changes in market conditions, or unforeseen events such as economic downturns.

Other Risks and Benefits:

The benefits of our machine learning application go beyond just salary prediction. It can also help companies with workforce planning and optimization by identifying skill gaps and predicting employee turnover. Additionally, it can help job seekers by providing insights into the job market and which industries and locations offer the best salary opportunities.

Other benefits of our model include its scalability and availability, as it can be deployed as a web application or API and can be accessed by businesses and organizations from anywhere in the world. Additionally, our model can be updated periodically to ensure its accuracy and reliability.

The risks associated with our model include the risk of data privacy and security, as the model requires access to sensitive employee data. It is important to implement appropriate security measures to protect the data and ensure compliance with data privacy regulations.

Conclusion

In this study, we evaluated the performance of three different classification models: Random Forest, SVC, and Decision Tree, to predict salary based on experience, age, job title, gender, and education.

Based on the evaluation, the Random Forest Classifier appears to be the most preferable model for the given dataset. The model achieved the highest accuracy score of 0.8901, which indicates that it can correctly predict the salary for approximately 89.01% of the data. Additionally, the Random Forest model achieved the lowest RMSE of 6833.98, which indicates that, on average, the predicted salary is approximately 6833.98 away from the actual salary. Moreover, the model achieved the highest R2 score of 0.9866, which suggests that the Random Forest model explains approximately 98.66% of the variance in the data.

These results indicate that the Random Forest model is capable of accurately predicting salary based on the available features.