1. \*\*Title and Overview:\*\*

- Provide a clear and concise title for your project.

- Write a brief overview that outlines the purpose, goals, and scope of the project.

**Adult Income Census Prediction**

This is a classification problem where we will predict whether the person gets salary >50K(greater than 50,000) or <=50K(less than or equal to 50,000). This is based on several factors such as age, workclass, education, martial-status, occupation, relationship, race, sex, capital-gain, capital-loss, hours per week, country, fnlwgt.

2. \*\*Introduction:\*\*

- Explain the problem or task that your machine learning project aims to solve.

This will let us know how a classification model can get trained to predict whether the given input gives >50K or <=50K label which will define the pay for a specified adult.

3. \*\*Objectives:\*\*

- Exploration data anlaysis

- Feature Engineering

- Machine Learning model training

- Prediction

4. \*\*Requirements:\*\*

- Jupyter notebook

- Python installation

- Pandas, Seaborn, sklearn, matplotlib, numpy

5. \*\*Data Collection:\*\*

- Gathered data from Kaggle.com

6. \*\*Data Exploration:\*\*

- Provide insights into the characteristics of the data, such as statistical summaries and visualizations.

7. \*\*Feature Engineering:\*\*

- Document the features used in your model.

- Explain any transformations or engineering applied to the features.

8. \*\*Model Architecture:\*\*

- Describe the architecture of your machine learning model.

- Include details about the types of layers, activation functions, and any other relevant information.

9. \*\*Training:\*\*

- Outline the training process, including hyperparameters and optimization algorithms.

- Provide information on how the model performance was evaluated during training.

10. \*\*Evaluation:\*\*

- Present the evaluation metrics used to assess the model's performance.

- Include any comparisons with baseline models or other relevant benchmarks.

11. \*\*Results:\*\*

- Showcase the results of your machine learning model, possibly through tables, charts, or visualizations.

- Discuss any insights gained from the results.

12. \*\*Challenges and Lessons Learned:\*\*

- Document any challenges faced during the project and how they were addressed.

- Share lessons learned and improvements for future work.

13. \*\*Conclusion:\*\*

- Summarize the key findings and outcomes of your machine learning project.

- Discuss the implications and potential applications.

14. \*\*Future Work:\*\*

- Suggest areas for future improvement or expansion of the project.

- Identify potential enhancements to the model or additional features.

15. \*\*References:\*\*

- Provide citations for any external datasets, papers, or resources used in your project.

16. \*\*Code Documentation:\*\*

- If your project involves code, make sure to provide detailed comments and documentation within the code itself. Consider using a consistent coding style.

17. \*\*Usage Instructions:\*\*

- Provide clear instructions on how to run your code and reproduce the results.

18. \*\*License:\*\*

- Specify the license under which your project is released.

Remember that documentation is an ongoing process. As your project evolves, update the documentation accordingly to keep it accurate and relevant.

The y-axis represents the number of people of a certain age, while the x-axis represents the age of each person. The blue line shows the number of persons earning less than or equal to $50,000, while the orange line represents those earning more than $50,000.

Both lines have a similar form, with a peak around the age of 30-35 and a slow fall as people age. However, the blue line is continuously higher than the orange line, suggesting that at each age value, more individuals earn less than or equal to $50,000 than persons who make more than $50,000.

We may conclude that there is a salary disparity between persons earning lower and higher wages and that this disparity remains throughout people's careers. Furthermore, the orange line peaked slightly later than the blue line, suggesting that persons earning more than $50,000 may take longer to attain their peak earning potential.