|  |  |
| --- | --- |
| Name Of The Student | Himanshu |
| Internship Project Topic | TCS iON RIO-210: Build a Classification Model for Drug Trials Dataset |
| Name of the Organization | TCS iON |
| Name of the Industry Mentor | Himdweep Walia |
| Name of the Institute | Amity University |

|  |  |  |
| --- | --- | --- |
| Date | Day # | Hours Spent |
| 07-06-2024 | Day-45 | 3.5 Hours |
| Activities done during the day:  **Project Hands-on – Conclusion**  **Link of the google drive google Colab file :-**  <https://colab.research.google.com/drive/1VQRq0l6oc9Uj4cOOqiuhkfS1JmpKr3fU?usp=sharing>  **Conclusion:** Best Parameters: {'n\_estimators': 200, 'min\_samples\_split': 2, 'min\_samples\_leaf': 1, 'max\_features': 'log2', 'max\_depth': None, 'bootstrap': True} Best Score: 0.83445625  In conclusion, after thorough exploration and evaluation of different classification algorithms, including KNeighbors Classifier and Random Forest Classifier, it was determined that both models yielded comparable performance metrics, with an average accuracy, precision, recall, and F1-score of approximately 83%.  While the Random Forest Classifier demonstrated strong predictive capabilities, the decision to select the KNeighbors Classifier was made based on various factors, including model interpretability, computational efficiency, and potential for fine-tuning in future iterations.  The KNeighbors Classifier, with its simplicity and ease of implementation, provides a transparent approach to classification tasks, making it easier to understand the underlying decision-making process. Additionally, its non-parametric nature allows for flexible adaptation to different types of data and problem domains.  Furthermore, the choice of the KNeighbors Classifier aligns with the principle of Occam's razor, favoring simpler models when achieving similar performance levels to more complex counterparts. This not only facilitates model comprehension but also reduces the risk of overfitting and improves generalization to unseen data.  Moving forward, the selected KNeighbors Classifier will serve as the foundation for further refinement and optimization, leveraging techniques such as hyperparameter tuning and feature engineering to enhance its predictive capabilities. Additionally, ongoing monitoring and evaluation will be conducted to ensure the model's effectiveness and reliability in real-world applications.  Overall, while acknowledging the competitive performance of the Random Forest Classifier, the decision to prioritize the KNeighbors Classifier reflects a balance between model performance, interpretability, and scalability, ultimately serving the objectives of the classification task effectively. | | |
|  | | |