**PRCP-1024-TexasSalaryPrediction**

**Problem Statement**

Task 1:-Prepare a complete data analysis report on the given data.

Task 2:-Create a predictive model which will help the Texas state government team to know the payroll information of employees of the state of Texas.

Task 3:-

* Who are the outliers in the salaries?
* What departments/roles have the biggest wage disparities between managers and employees?
* Have salaries and total compensations for some roles/ departments/ head-count changed over time?

**Dataset Link:**

This database has salary information for positions at all 113 agencies in the Texas state government. The Tribune obtained this data by requesting salary records from the state comptroller, as allowed by the Texas Public Information Act.

Link : <https://d3ilbtxij3aepc.cloudfront.net/projects/CDS-Capstone-Projects/salary.zip>

**Attribute Information :**

Agency

Agency Name

Last Name

First Name

MI (Middle Initial)

Class title

Ethnicity

Gender

Status

Employ Date

Hourly rate

Hrs per week

Monthly (Monthly income)

Annual (Annual Income)

State number

**Model Comparison Report**

Create a report stating the performance of multiple models on this data and suggest the best model for production.

**Report on Challenges faced**

Create a report which should include challenges you faced on data and what technique used with proper reason.

Note:-All above task has to be created on single jupyter notebook and share the same while final submission of project.