NAAN MUDHALVAN DATASCIENCE FUDAMANETAL PROJECT

PROJECT TITLE:
DATA SCIENCE JOB ANALYSIS

PRESENTED BY:
VETRIKIANNAN.K
SACS M.A.V.M.M. ENGINEERING COLLEGE
III YR CIVIL DEPT

Project statement

- Our goal is to complete the task below based off the article and see if you reach a similar conclusion. determine the job and salary analysis.

Problem solution

- By Utilizing the pandas and visualization skills to data science job analysis
- The movie rating will be separated and explored by the given csv datasets.

System approach

- Building the proposed solution would involve a combination of data processing, pyth0n programming and visualization skills.
- **System requirements:**
  1. **Hardware** :
     - A computer with sufficient processing power, preferably with multiple cores or a GPU for faster training of machine learning models.
     - Adequate RAM to handle the size of the dataset and computational requirements.
  2. **Software:**
     - - An operating system compatible with the required python libraries (e.g., windows,linux,macOS).

System approach – CONT.

- Library Requirements:
- Data processing and analysis:
  - **Pandas :** For data manipulation and analysis.
  - **Numpy :** For numerical operations on data.
- Data visualization:
  - **Matplotlib and seaborn**: For creating visualizations to understand data patterns.
  - **Plotly or Bokeh**: Interactive visualization libraries for more complex visualizations.

Algorithm selection

Data exploration :

- Explore the movie rating structure, features and variables.
- Identify potential patterns, correlations and outliners.

Problem formulation:

- Define the problem: Predict optimal booking times, likelihood of special requests based on historical data.

Algorithm selection:

- Regression tasks(e.g., predicting daily rates)
  - Consider linear regression, decision tree, or ensemble methods
- classification tasks(e.g., predicting special requests);
  - Consider logistic regressive, decision trees or random forests.

ALGORITHM & DEPLOYMENT

Data input:

Data collection:

- Gather historical data including booking dates ,special requests,and relevant details.

Data cleaning:

- Handle missing values, outliners, and any inconsistencies in the dataset.
- Convert categorical variables into numerical representation through encoding techniques.

Feature Engineering:

- Create new features or modify existing ones based on domain knowledge.
- Extract meaningful information from date variables, such as day-of-week or month.

Training process:

Data splitting:

- Divide the dataset into training and testing sets to evaluate the model's performance.

Feature scaling:

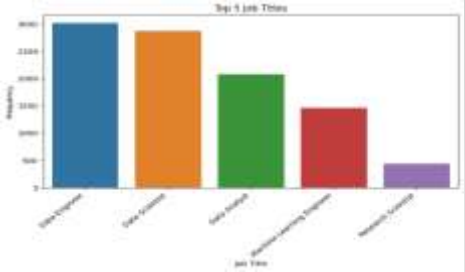
- Standardize or normalize numerical features to ensure they have consistent scale.

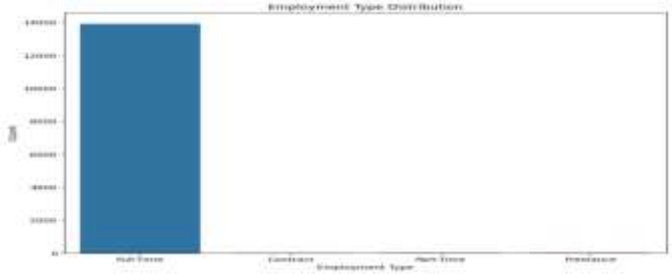
Modeling training:

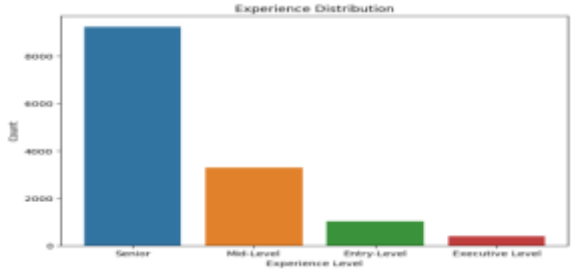
- Use the selected algorithm to train the model on the all sites scores dataset.
- Adjust hyperparameters to optimize model performance.

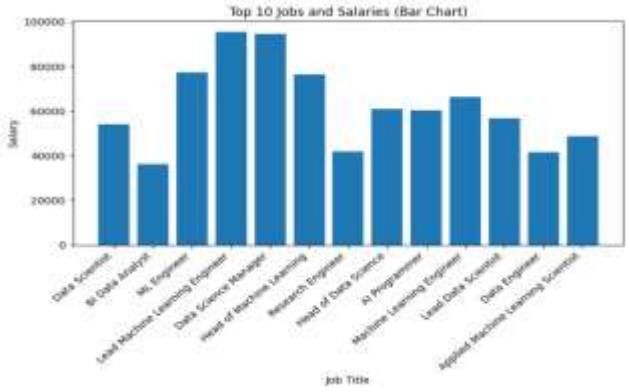
Model evaluation:

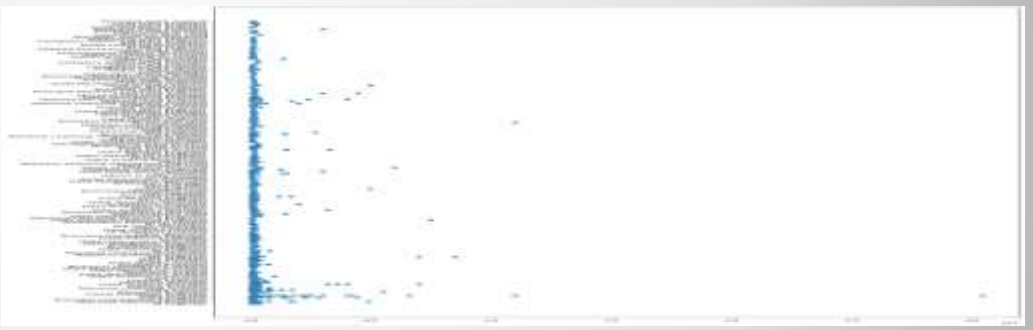
- Evaluate the model on the dataset using appropriate metrics(e.g., Mean Squared Error for regression,accuracy,precision, recall for classification).
- Fine-tune the model if necessary.

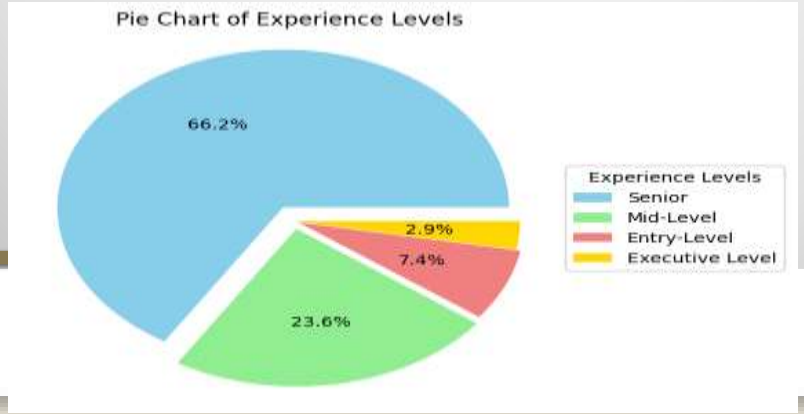

Top 5 Job Titles


Employment Type Distribution


Experience Distribution
Senior
Mid-Level
Entry-Level
Executive Level
Experience Level


Top 10 Jobs and Salaries (Bar Chart)
Salary
Job Title
Data Scientist
BI Data Analyst
ML Engineer
Lead Machine Learning Engineer
Data Science Manager
Head of Machine Learning
Research Engineer
Head of Data Science
AI Programmer
Machine Learning Engineer
Lead Data Scientist
Data Engineer
Applied Machine Learning Scientist




Pie Chart of Experience Levels
66.2%
23.6%
7.4%
2.9%
Experience Levels
Senior
Mid-Level
Entry-Level
Executive Level

RESUL

CONCLUSION

- It proved that and we have data science job analysis was done and analyze fully by visualstudio

REFERENCES

- http://www.kaggle.com/datasets
- http://pandas.pydata.org/pandas-docs/stable/user guide/ index.html
- http://seaborn.pydata.org/
- http://matplotlib.org/stable/contents.html

Thank you