Project Title	**Predicting Customer Behavior in DVD Rental Using Deep Learning with AWS deployment**
Skills take away From This Project	Data Cleaning and Preprocessing Exploratory Data Analysis (EDA) Feature Engineering for Neural Networks Building and Training Artificial Neural Networks (ANNs) Model Evaluation and Optimization postgresql Deploying Deep Learning Models using AWS and Streamlit Documentation and Reporting
Domain	**Entertainment / Rental Industry**

Problem Statement:

Predict customer behavior in a DVD rental business, specifically focusing on customer churn, movie genre preferences, and rental demand forecasting. The objective is to use deep learning techniques to create models that can assist in decision-making, marketing strategies, and personalized customer recommendations.

Business Use Cases:

- **Customer Churn Prediction**: Identify customers likely to churn so proactive retention strategies can be applied.

- **Movie Genre Preference Prediction**: Recommend personalized movie genres to customers based on past rental behavior.

- **Rental Demand Forecasting**: Predict future rental demand to optimize stock and inventory planning at DVD rental stores.

Approach:

- **Data Understanding & Preprocessing**:

- Collect and preprocess data from the DVD rental dataset.
- Handle missing values, normalize numerical data, and encode categorical data.

🎬 **Exploratory Data Analysis (EDA)**:

- Conduct data exploration to understand the relationships between variables, using visualizations like histograms, box plots, and correlation matrices.

🎬 **Feature Engineering**:

- Generate useful features from customer demographics, rental history, movie genres, and other transactional data.

🎬 **Model Development**:

- Build an Artificial Neural Network (ANN) to predict customer churn (classification task).
- Build a multi-class classification model for genre prediction.
- Use RNN/LSTM models for rental demand forecasting (time-series analysis).

🎬 **Model Evaluation**:

- Evaluate models using appropriate metrics such as accuracy, precision, recall, F1-score, and RMSE for regression tasks.

🎬 **Deployment**:

- Deploy the best-performing model using AWS and integrate it with a Streamlit application for interactive use.

Results:

🎬 Predict customer churn with an accuracy of over 85%.

🎬 Predict customer movie genre preferences with a multi-class classification accuracy of 80%.

🎬 Forecast DVD rental demand with a mean absolute error (MAE) of less than 10 rentals per store.

🎬 Provide actionable insights for reducing customer churn and enhancing inventory management.

Project Evaluation metrics:

🎬 **For Churn Prediction:**

- **Accuracy, Precision, Recall, F1-Score**

🎬 **For Genre Prediction:**

- **Accuracy, Confusion Matrix, Precision, Recall, F1-Score**

🎬 **For Demand Forecasting:**

- **MAE (Mean Absolute Error), RMSE (Root Mean Squared Error)**

Technical Tags:

🎬 Data Cleaning

🎬 Feature Engineering

🎬 Artificial Neural Networks (ANN)

🎬 Postgresql

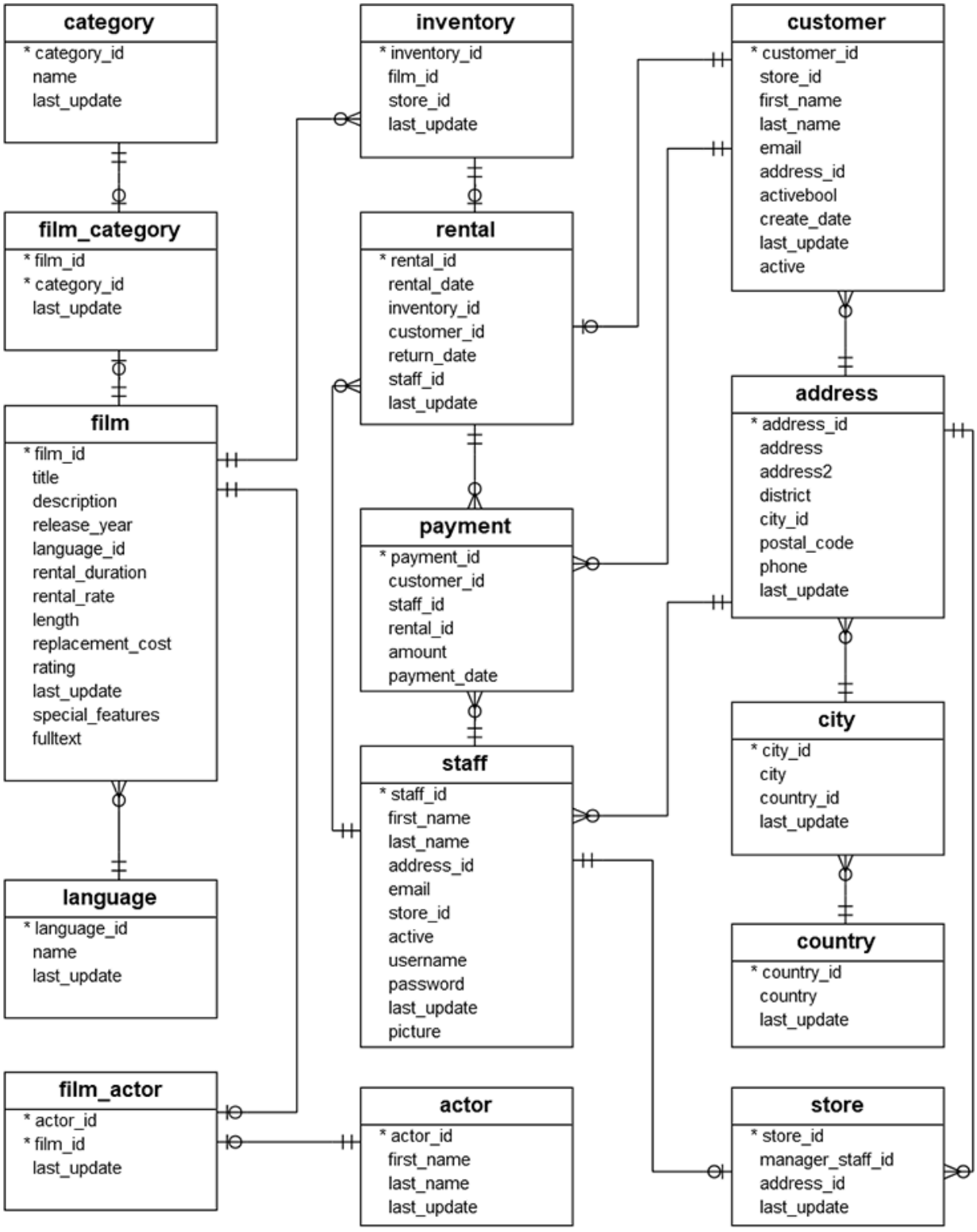
🎬 Churn Prediction

🎬 Time-Series Forecasting

🎬 AWS Deployment

🎬 Streamlit for Frontend

Data Set:



Data Set Explanation:

1. Actor Table

- Purpose: Contains information about the actors who have appeared in the films available for rental.

- Key Columns:

  - actor_id: Unique identifier for each actor.

  - first_name and last_name: Name of the actor.

  - last_update: Timestamp indicating when the record was last updated.

2. Film Table

- Purpose: Contains details about the films available for rental in the store.

- Key Columns:

  - film_id: Unique identifier for each film.

  - title: Title of the film.

  - description: Short description of the film.

  - release_year: Year the film was released.

  - rental_duration: Number of days the film can be rented.

  - rental_rate: Rental cost of the film.

  - rating: Film rating (e.g., PG, R).

3. Film Actor Table

- Purpose: A linking table that establishes the many-to-many relationship between films and actors, showing which actors appeared in which films.

- Key Columns:

  - actor_id: References the actor involved in the film.

  - film_id: References the film the actor appeared in.

4. Film Category Table

- Purpose: Links each film to its category (genre), allowing for categorization of films based on genre.

- Key Columns:

  - film_id: References the film being categorized.

  - category_id: References the genre (e.g., Action, Drama) of the film.

5. Category Table

- Purpose: Stores the available categories (genres) for films.

- Key Columns:

  - category_id: Unique identifier for each category.

  - name: Name of the category (e.g., Action, Comedy).

  - last_update: Timestamp indicating the last update of the record.

6. Customer Table

- Purpose: Contains personal and account information of the customers renting DVDs.

- Key Columns:

  - customer_id: Unique identifier for each customer.

  - first_name and last_name: Name of the customer.

  - email: Customer's email address.

  - active: Indicates whether the customer is currently active.

  - store_id: References the store the customer is associated with.

7. Rental Table

- Purpose: Stores transactional data about each DVD rental, including rental and return dates.

- Key Columns:

  - rental_id: Unique identifier for each rental transaction.

  - rental_date: The date when the DVD was rented.

  - return_date: The date when the DVD was returned.

  - customer_id: References the customer who rented the DVD.

  - inventory_id: References the specific DVD copy that was rented.

8. Inventory Table

- Purpose: Stores information about individual copies of each film available for rent at different store locations.

- Key Columns:

  - inventory_id: Unique identifier for each copy of a film.

  - film_id: References the specific film.

  - store_id: References the store where the DVD is available.

9. Payment Table

- Purpose: Tracks payments made by customers for their rentals.

- Key Columns:

  - payment_id: Unique identifier for each payment transaction.

  - customer_id: References the customer making the payment.

  - amount: The amount paid by the customer.

- payment_date: The date when the payment was made.

10. Store Table

- Purpose: Contains information about the store locations where films can be rented.

- Key Columns:

  - store_id: Unique identifier for each store.

  - manager_staff_id: References the staff member managing the store.

  - address_id: References the address of the store.

11. Staff Table

- Purpose: Contains details of the employees working in the DVD rental stores.

- Key Columns:

  - staff_id: Unique identifier for each staff member.

  - first_name and last_name: Name of the staff member.

  - email: Staff member's email address.

  - store_id: References the store where the staff member works.

12. Language Table

- Purpose: Stores information about the languages available for the films in the rental store.

- Key Columns:

  - language_id: Unique identifier for each language.

  - name: Name of the language (e.g., English, French).

Project Deliverables:

- Cleaned and preprocessed dataset

- EDA report with visualizations (optional)

- Feature engineering code and descriptions

- Predictive models with code and explanations

- Model evaluation report

- Insights and recommendations report

- AWS Deployment with nohup

- Source code and documentation

Project Guidelines:

- Follow coding standards and best practices (PEP 8 for Python).

- Use version control (e.g., Git) to manage code.

- Document all steps clearly, including data cleaning, feature engineering, modeling, and evaluation.

- Ensure reproducibility of results.

Timeline:

1. **Week 1**: Data Cleaning and Preparation

2. **Week 1**: Exploratory Data Analysis (EDA) (optional)

3. **Week 1**: Feature Engineering

4. **Week 2**: ANN Modeling and Evaluation

5. **Week 2**: AWS deployment using streamlit and Recommendations

6. **Week 2**: Final Report and Submission

Milestones:

- End of Week 1: Cleaned dataset ready

- End of Week 1: EDA report completed

- End of Week 1: Feature engineering completed

- End of Week 2: Models developed and evaluated

- End of Week 2: Insights and recommendations finalized

- End of Week 2: Final submission

PROJECT DOUBT CLARIFICATION SESSION (PROJECT AND CLASS DOUBTS)

About Session: The Project Doubt Clarification Session is a helpful resource for resolving questions and concerns about projects and class topics. It provides support in understanding project requirements, addressing code issues, and clarifying class concepts. The session aims to enhance comprehension and provide guidance to overcome challenges effectively.
Note: Book the slot at least before 12:00 Pm on the same day

Timing: Tuesday, Thursday, Saturday (5:00PM to 7:00PM)

Booking link :https://forms.gle/XC553oSbMJ2Gcfug9

LIVE EVALUATION SESSION (CAPSTONE AND FINAL PROJECT)

About Session: The Live Evaluation Session for Capstone and Final Projects allows participants to showcase their projects and receive real-time feedback for improvement. It assesses project quality and provides an opportunity for discussion and evaluation.
Note: This form will Open on Saturday and Sunday Only on Every Week

Timing: Monday-Saturday (11:30PM to 12:30PM)

Booking link : https://forms.gle/1m2Gsro41fLtZurRA