Assignment 3: Telco Customer Churn Prediction

for AIDI1009-25S: Neural Networks

—

Due Date: **2ⁿᵈ August 2025 @11:55 PM**

Customer churn prediction is a common real-world problem in various industries, particularly in telecommunications, e-commerce, and finance. This assignment aims to develop a deep learning model that can predict which customers are likely to stop using a product or service, allowing businesses to take proactive steps to retain them.to retain customers, improve satisfaction, and increase revenue.

Objective

Develop a deep learning model to predict customer churn based on historical customer data. You will explore various two classification techniques, evaluate model performance, and provide actionable insights based on the features that contribute most to customer churn.

Dataset

1. **Dataset Source:** You can use a publicly available dataset like the Telco Customer Churn dataset from Kaggle https://www.kaggle.com/datasets/blastchar/telco-customer-churn or a similar dataset with customer information and churn status.

2. **Features:** Common features include age, tenure (how long a customer has been with the company), payment type, contract type, total charges, number of products used, and whether the customer is subscribed to additional services.

3. **Target:** The target variable is a binary variable indicating whether the customer churned (1) or not (0).

1 Assignment Phases

1.1 Data Exploration and Preprocessing

1. Load and explore the dataset, understanding the features and their distributions..

2. Clean the data by handling missing values and converting categorical variables into numerical representations (one-hot encoding, label encoding).

3. Perform exploratory data analysis (EDA) to identify patterns and relationships between features and churn.

4. Visualize key findings using histograms, scatter plots, etc.

1.2 Feature Engineering and Scaling

1. **Feature Engineering:** Create meaningful features that may help the model, such as calculating average monthly charges or segmenting customers by contract type.

2. **Encoding:** Convert categorical variables into numerical formats.

3. **Scaling:** Normalize or scale features as needed, especially when working with models sensitive to feature magnitude.

1.3 Model Building

1. **Train-Test Split:** Split the data into training and testing sets to evaluate model performance.

2. **Model-1:** You have to implement Neural networks (using Tensorflow/Pytorch/Keras) on this dataset.

3. **Model-2:** You have to implement Convolutional Neural networks (CNNs) (using Tensorflow/Pytorch/Keras) on this dataset.

4. Justify the choice of model based on the characteristics of the dataset and problem (i.e., which one performed better and why).

1.4 Model Training and Evaluation

1. Split the data into training and testing sets.

2. Train the selected model(s) on the training data.

3. Evaluate model performance on the testing data using metrics like accuracy, precision, recall, F1-score, and AUC (Area Under the ROC Curve).

4. **Deployment (Optional):** If possible, deploy the trained model as a simple web application or API to demonstrate real-world applicability.

2 Deliverables

2.1 Assignment Report:

1. Summarize data preprocessing, feature engineering, model results, and key findings.

2. Describe the impact of different features on churn prediction and give recommendations for customer retention based on model insights.

2.2 Codebase:

- A well-organized codebase, ideally in Jupyter notebooks, with comments explaining each step of the process.

3 Report Format

Name your report AIDI1009-24F-10827: Assignment#3 LastNameFirstName.pdf. Below is the general format of the report required:

1. The front page (i.e. title page) should contain only the following:

   - Course #, Course Name and Date
   - Your Name and ID
   - Assignment # and title of the assignment.

2. Introduce the problem to be solved:

   - Problem Statement
     (a) Briefly describe the problem solved in the assignment.
     (b) Assumptions and Constraints
     (c) Constraints could be for example using certain libraries, datasets, or a specific programming language.

3. Answer all questions posed in Section 3. Append the following to your answers:

   (a) Plots and Graphs.
   (b) Tables.

    (c) Figures.

4. Python Code

    (a) Python code for the problem

    (b) Document your code.