Submitted by :

Nimra DHC-458

AI/ML Engineering Internship – Task Reports

✅ Task 1: News Topic Classifier Using BERT

Objective: Fine-tune `bert-base-uncased` model on AG News dataset to classify news into categories.

Tools Used: Hugging Face Transformers, Datasets, PyTorch

Steps:
- Loaded AG News dataset
- Tokenized using BERT tokenizer
- Fine-tuned `bert-base-uncased` with Trainer API
- Evaluated using accuracy and F1-score

Results:
- Accuracy and F1-score achieved on validation subset (sample: 100 train, 50 eval)

Key Learnings:
- Transfer learning using BERT
- Text classification with Transformers
- Using Trainer API and tokenization

✅ Task 2: End-to-End ML Pipeline for Customer Churn

Objective: Build a reusable ML pipeline to predict customer churn using Telco dataset.

Tools Used: Scikit-learn, Pandas, Joblib

Steps:

- Loaded and cleaned Telco churn data
- Preprocessed with `ColumnTransformer` (scaling + encoding)
- Built pipeline using `Pipeline()`
- Trained models: Logistic Regression & Random Forest
- Used `GridSearchCV` for hyperparameter tuning
- Exported final pipeline using `joblib`

Results:
- Achieved accuracy and F1-score using pipeline

Key Learnings:
- ML pipeline structuring
- Feature encoding and scaling
- Grid search for tuning
- Exporting models for reuse

✅ Task 5: Auto Tagging Support Tickets Using LLM

Objective: Automatically classify support ticket text into tags using zero-shot classification.

Tools Used: Hugging Face Transformers, `facebook/bart-large-mnli`

Steps:
- Loaded pre-trained model using pipeline API
- Defined ticket and candidate labels (tags)
- Used zero-shot classification (no training needed)
- Output top 3 predicted tags with scores

Results:
- Tickets correctly matched to categories with high confidence

Key Learnings:
- Zero-shot classification
- LLM-based text tagging
- No-training inference via Hugging Face pipeline