Twitter Bot Detection using Cresci-2017

Project Objective:

Replicate an existing BERT + metadata bot detection model and later improve it by replacing BERT with lightweight transformers.

Dataset:

Cresci-2017 (Twitter-only bot detection dataset)

Files Used:

node – user metadata and text

label – bot/human labels

split – train/validation/test split

Files Ignored:

edge, user_info.pt

Data Preparation:

Each Twitter account is one sample.

Combine bio and recent tweets as text.

Extract basic metadata (followers, following, tweets, account age, verified).

Assign labels using label file.

Use predefined splits.

Baseline Architecture:

Text branch: BERT-base (frozen)

Metadata branch: Simple MLP

Fusion: Concatenation

Classifier: Dense layers + Sigmoid

Training:

Train classifier and metadata layers only.

Use binary classification loss.

Evaluate using Accuracy and F1-score.

Baseline Result:

Create a table showing BERT + Metadata performance.

Extension:

Replace BERT with DistilBERT or MiniLM.

Keep everything else unchanged.

Compare accuracy, F1-score, and efficiency.

Viva Line:

First, we replicate an existing BERT-based hybrid bot detection model on Cresci-2017, then improve it using lightweight transformers.