

In-Context Learning

The following task focuses on evaluating whether a Twitter user is a bot or human with the help of several labeled examples.

<in-context examples>

Label: bot

Username: <username> Follower count: 309 Following count: 1412 Tweet count: 1745 Verified: False Active years: 12 years

Description

The following task focuses on evaluating whether a Twitter user is a bot or human with the help of the user's self-written description.

<in-context examples>

Description: A marketer in and out. Writes on marketing & sometimes straight from the heart Check out at <link> Label: bot

Structure 200

The following task focuses on evaluating whether a Twitter user is a bot or human with the help of the user's followers and followings and their labels.

These users follow the target

The target user follows these users: <user information> Target user: <target user information> Label: human

user: <user information>

Instruction Tuning (3)



Textual Manipulation

Zero-shot

Few-shot

Classifier Guide

Below is a description of a Twitter user and its variants, paired with their score predicted by a bot classifier. Please rewrite the description to make the user appear more human.

Description: <description at step 0> Score: 0.68

Description: <description at step i> Score: 0 26 Description: <description at step i+1>

follow (1-k):

Below is a target Twitter bot and five potential new users to

Add Neighbor

follow. Please suggest one new user to follow so that the target bot appears more human Target Bot:

<description and metadata> Potential Followings: user 0: <description and metadata>

Please select one user to

Remove Neighbor Below is a target Twitter bot and five potential users to

unfollow. Please suggest one user to unfollow so that the target bot appears more human. Target Bot:

<description and metadata> Potential users to unfollow:

Structural Manipulation

user 0: <description and metadata>

Please select one user to unfollow (1-k):

> (j?)

Both (add & remove)

Text Attribute

Step 1:

Bot Descriptions: <top-k similar bot descriptions> Human Descriptions: <top-k similar human descriptions> Compare and give the key distinct feature of human's descriptions: <generated attribute>

Step 2:

<generated attribute>

Based on the description, paraphrase this to human description:

Bot: <description> Human:

(a) opportunities

(b) risks