



Prompts

Task-oriented

Role-oriented

Zero-shot

Few-shot



Attacks

Character-level

DeepWordBug TextBugger

Word-level

TextFooler BertAttack

Sentence-level

CheckList StressTest

Semantic-level

Human-crafted



Models

Flan-T5-large (0.8B)

Dolly-6B

Cerebras-13B

LLaMA-13B

Vicuna-13B

GPT-NEOX-20B

Flan-UL2 (20B)

ChatGPT



Tasks

Sentiment Analysis

Grammar Correctness

Duplicate Sentence Detection

Natural Language Inference

Multi-task Knowledge

Reading Comprehension

Translation

Math

<u>ılı.</u>

Datasets

GLUE

MMLU

SQuAD V2

UN Multi

IWSLT 2017

Mathematics