

# Natural Language Processing: Minimizing Bias and Misunderstanding for AI Models in Understanding and Generating Human-like Text

MSc Research Project  
Master of Science in Artificial Intelligence

Chetna Verma  
Student ID: x22168842

School of Computing  
National College of Ireland

Supervisor: Mayank Jain

National College of Ireland  
Project Submission Sheet  
School of Computing



<b>Student Name:</b>	Chetna Verma
<b>Student ID:</b>	x22168842
<b>Programme:</b>	Master of Science in Artificial Intelligence
<b>Year:</b>	2023 January - 2024 January
<b>Module:</b>	MSc Research Project
<b>Supervisor:</b>	Mayank Jain
<b>Submission Due Date:</b>	31/01/2024
<b>Project Title:</b>	Natural Language Processing: Minimizing Bias and Misunderstanding for AI Models in Understanding and Generating Human-like Text
<b>Word Count:</b>	518
<b>Page Count:</b>	3

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

**ALL** internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

<b>Signature:</b>	Chetna
<b>Date:</b>	31st January 2024

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:**

Attach a completed copy of this sheet to each project (including multiple copies).	✓
<b>Attach a Moodle submission receipt of the online project submission</b> , to each project (including multiple copies).	✓
<b>You must ensure that you retain a HARD COPY of the project</b> , both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

<b>Office Use Only</b>	
Signature:	
Date:	
Penalty Applied (if applicable):	

# Natural Language Processing: Minimizing Bias and Misunderstanding for AI Models in Understanding and Generating Human-like Text

Chetna Verma  
x22168842

## 1 System Requirements

Operating System:

Linux (Ubuntu 18.04 or later), macOS (10.13 or later), or Windows 10. Ensure the OS is updated and has all necessary drivers installed, especially for GPU support.

Python Environment:

Python 3.6 or higher. It's recommended to use a virtual environment (like `venv` or `conda`) for Python to manage dependencies effectively. *Download PyCharm: Python IDE for Professional Developers by JetBrains* (2021)

Hardware Requirements:

CPU: A modern multi-core CPU (e.g., Intel Core i7/i9, AMD Ryzen 7/9) for efficient processing.

GPU: A dedicated NVIDIA GPU with CUDA support. *CUDA Toolkit Documentation 12.3 Update 1* (n.d.) is highly recommended for faster training and inference. GPUs like NVIDIA RTX 2070, 2080, 3080, or equivalent are suitable.

RAM: Minimum 16 GB RAM; 32 GB or higher is preferable for handling large datasets and intensive computations.

Storage: SSD (Solid State Drive) with at least 50 GB of free space for faster data read/write operations. Additional space for model checkpoints and data storage.

Software Dependencies:

PyTorch: Ensure that a compatible version of PyTorch is installed for GPU/CPU. Install the CUDA toolkit compatible with your PyTorch and GPU version if using a GPU. *NVIDIA CUDA Toolkit 12.1 downloads* (n.d.)

Transformers Library: The Hugging Face transformers library for loading and using GPT-2 models. *Transformers* (n.d.)

NLTK Library: For natural language processing tasks and tokenization. *NLTK :: Natural Language Toolkit* (n.d.)

Other Python libraries such as numpy, pandas, for data manipulation and matplotlib for visualization might be useful.

## 2 Contextual Token Generation

Prerequisites:

Python environment with NLTK library.

corpus.txt file with original text data.

NLTK Setup: Ensure NLTK is installed and download the necessary components:

```
nltk.download('punkt')
nltk.download('stopwords')
```

Token Generation: Run the script for token generation, which preprocesses the text, generates contextual tokens, and modifies the original corpus.

Outputs: new\_corpus.txt and new\_tokens.txt are generated as outputs.

## 3 Training the Enhanced GPT-2 Model Environment Setup

PyTorch environment with GPU support (if available). Wolf (2019)

Install the "transformers" library from Hugging Face.

Model Training:

Load the GPT-2 model and tokenizer.

Update the tokenizer with new tokens from new\_tokens.txt.

Load the modified corpus new\_corpus.txt for training.

Set training parameters like batch size, learning rate, and number of epochs.

Train the model, and monitor loss during training for each epoch.

Save the trained model and tokenizer to MODEL\_DIRECTORY.

## 4 Prompting with Trained Model

Model Loading:

Load the trained model and tokenizer from MODEL\_DIRECTORY.

Prompting Function:

Create a function `generate_text` for generating text based on a given prompt.

The function takes a prompt, generates text using the model, and decodes the output to a human-readable format.

Ensure the model is in evaluation mode and utilize GPU if available.

Usage:

Run the `generate_text` function with a test prompt to generate text.

## 5 Additional Notes

Ensure the paths for loading models, tokenizers, and data are correctly set.

Adjust the window size and similarity threshold in the token generation script according to your specific requirements.

Monitor the GPU utilization and adjust the batch size during training if necessary to optimise performance.

Regularly save model checkpoints during training to prevent data loss.

## References

*CUDA Toolkit Documentation 12.3 Update 1* (n.d.).

**URL:** <https://docs.nvidia.com/cuda/index.html>

*Download PyCharm: Python IDE for Professional Developers by JetBrains* (2021).

**URL:** <https://www.jetbrains.com/pycharm/download/?section=windows>

*NLTK :: Natural Language Toolkit* (n.d.).

**URL:** <https://www.nltk.org/>

*NVIDIA CUDA Toolkit 12.1 downloads* (n.d.).

**URL:** <https://developer.nvidia.com/cuda-downloads>

Wolf, T. (2019). HuggingFace’s Transformers: State-of-the-art natural language processing.

**URL:** <https://arxiv.org/abs/1910.03771>

*Transformers* (n.d.).

**URL:** <https://huggingface.co/transformers/>