**MSc Research Project – Custom Dataset and Source Code Links**

NSPCC Custom Dataset

<https://huggingface.co/datasets/scott156/NSPCC_dataset>

Google Colab Notebooks

|  |  |
| --- | --- |
| **Notebook Title & Link** | **Description** |
| [Train - LongT5Base.ipynb](https://colab.research.google.com/drive/13XPZBiqdjODCoHjD4WkX-V8NqF85So-O?usp=sharing) | Contains the code which utilised full fine-tuning on the LongT5 base model using the NSPCC custom dataset. |
| [Train - LongT5Large.ipynb](https://colab.research.google.com/drive/1qkfehHi4pOb98MA5tBdA9Yy9TW09KOCW?usp=sharing) | Contains the code which utilised full fine-tuning on the LongT5 large model using the NSPCC custom dataset. |
| [Train - LongT5XL.ipynb](https://colab.research.google.com/drive/1WLl3QVGGZ-CiJqH3Fr2d16cPT04Jem2R?usp=sharing) | Contains the code which utilised Parameter Efficient Fine-Tuning (PEFT) on the LongT5 XL model using the NSPCC custom dataset. The PEFT method adopted was Quantized Low-Rank Adaptation (QLoRA). |
| [Train - LEDBase.ipynb](https://colab.research.google.com/drive/1CjqJN7JSnpKOChez9Kd84yI7Xna33Gud?usp=sharing) | Contains the code which utilised full fine-tuning on the LED base model using the NSPCC custom dataset. |
| [Train - LEDLarge.ipynb](https://colab.research.google.com/drive/11sqGN7lLSIXFJmailODv4WmTODxGxgSm?usp=sharing) | Contains the code which utilised full fine-tuning on the LED large model using the NSPCC custom dataset. |
| [Test - Generate Summaries.ipynb](https://colab.research.google.com/drive/1rg39TL_gdFq87yLiAhtWHtfjNg4B3HbP?usp=sharing) | Contains the code which generated summaries of the reports in the test set using each fine-tuned model. |
| [Test - Calculate Metrics.ipynb](https://colab.research.google.com/drive/19uw7QUEL4qj_0jXFhHSuyM1_z7kmqA3P?usp=sharing) | Uses the summaries generated from the code above to calculate ROUGE and BERTScore metrics for each fine-tuned model. |
| [Test - Create Comparison Table.ipynb](https://colab.research.google.com/drive/1pmt2C1uVXHKT2FXy4F3WG_VLGpOD7PYx?usp=sharing) | Creates a table that allows a comparison of summaries across all fine-tuned models. |