AI for the Arts & Humanities (A) Lab Book

Weeks 5 - 11: Deep Dive Machine Learning

**This is the last lab sheet for this course.**

Many of the tasks in this document and associated notebooks will require you to discuss with your group to address them. You will be expected **not only to attempt some tasks yourself, but to explain your successes and difficulties to others in your group** to make progress going forward.

While doing each of the Parts and discussing with others, ask the following questions to yourself:

* Are there analogies you can use to explain your problems or solutions?
* Can you break down the task to smaller topics areas to explain your problems or solutions?
* Are there resources that can help with any part of the task or to add relevant context?

Use these to enhance your discussions and content of your notebooks. People have different backgrounds. By attempting to explain, and by watching others explain, you create an environment for better understanding for yourself and for those in your group. The article [“Embracing Digitalization: Student Learning and New Technologies”](https://journals.sagepub.com/doi/10.1177/0273475318820895) (Crittenden, Biel & Lovely 2018) demonstrates how we learn and retain more information when we explain and discuss it with others.

**This will also help you better “present your code in your notebook” to engage the broader arts and humanities audience – one of the intended learning outcomes of the course and your assessment.**

# Part 1: Deep Dive Machine Learning (Group Task)

The tasks in this part of the exercise should be carried out with your peer group!

**You are required to do the tasks in Part 1 with your peer group!**

## Your Task:

Look through the notebook designated below for this part of the lab. This notebook takes you through the machine learning pipeline using two examples. In the notebook, you will find code, textual explanations, explanations in the comments. You will also find questions throughout the notebook accompanied by pre-created markdown cells for you to modify for your own text to answer the questions.

Your task is to explore and run the code, and address the questions related to the machine learning process in discussion with your peer group.

## Designated Notebook

In the [Lab Resources](https://moodle.gla.ac.uk/course/view.php?id=39566#section-3) section of the Moodle, you will find a Jupyter Notebook [“**Machine-Learning-by-Example-from-Start-to-End”**](https://moodle.gla.ac.uk/mod/resource/view.php?id=4095655)**.**

## Workflow You Should Follow:

* Download the designated notebook for Part 1.
* Open Anaconda, launch Jupyter Notebook, and open the downloaded notebook and follow the instructions therein.
* Before you finish your lab, upload your notebook to your GitHub repository – even if you have not finished it.
* It is good practice to create multiple commit points, for example, after each **Task** or **Part** is completed with an appropriate message to record progress.
* Do not wait until all the lab notebooks and exercises are all ready and complete. This is not good information management practice and misses out on some of the advantages of using GitHub.
* This helps you keep track of what needs further work and allows you to work on different parts in parallel.

# Part 2: Critically Engaging with AI Ethics

A core part of responsible AI research relates to thinking about ethics. This is a ongoing task, throughout your machine learning pipeline and further, from data collection to deployment in the real world.

## Your Task:

Look through the notebook designated below for this part of the lab. This notebook introduces you to tutorials on Kaggle for [AI Ethics](https://www.kaggle.com/learn/intro-to-ai-ethics) and [Machine Learning Explainability](https://www.kaggle.com/learn/machine-learning-explainability) to help you understand some of the key concepts in AI ethics: **bias, fairness, and explainability**.

You will be expected to work through parts of the Kaggle tutorials while taking notes, and screenshots of selected parts. You will use these to discuss the questions with your group, and report the results in a markdown cell, pre-created for you to modify.

## Designated Notebook

In the [Lab Resources](https://moodle.gla.ac.uk/course/view.php?id=39566#section-3) section, you will find a Jupyter Notebook linked at [“**Critically Engaging With AI Ethics”**](https://moodle.gla.ac.uk/mod/resource/view.php?id=4095656)**.**

## Workflow You Should Follow:

* Download the designated notebook for Part 2.
* Open Anaconda, launch Jupyter Notebook, and open the downloaded notebook and follow the instructions therein.
* Before you finish your lab, upload your notebook to your GitHub repository – even if you have not finished it.
* It is good practice to create multiple commit points, for example, after each **Task** or **Part** is completed with an appropriate message to record progress.
* Do not wait until all the lab notebooks and exercises are all ready and complete. This is not good information management practice and misses out on some of the advantages of using GitHub.

# Part 3: Understanding and Presenting your First Generative AI Code

In this section you will work with generative AI models as an opportunity to contemplate Human-AI creativity, a hot topic in recent years.

## Your Task

You have the option of choosing one the designated notebooks below: a notebook that generates music, or one that generates text.

You will work with one of the two designated notebooks below to present the code to the arts and humanities using your notebook, adhering to the portfolio assessment criteria, including:

1. You engagement with the accuracy of the presented information (20%): for example, using self-assessment, and verification of the results of running relevant code, and/or other displayed content. You could also make explicit peer group discussions related to the tasks, as well as other examples, and sources you used to verify correctness.
2. Explanations and context (20%): for example, demonstrating the use of markup sections to explain the broader context of your portfolio, tasks, and/or objectives. You might contrast the expected results of your code, and the results. Perhaps employ commenting blocks and lines of code to explain intention, where it is deemed that your target audience would require it. Peer group discussion can help here also.
3. Selection of informative names (20%): e.g. for variables, functions, objects, and files, to help you and your target audience make sense of your code (avoid names that make sense only to yourself at the time).
4. Exploration of ways you might enhance user engagement and experience with your code (20%): for example, by including displays of examples, multimedia and/or widgets.
5. Attention to integrity (20%): for example, be mindful of whether or not you have permission to use the information in your portfolio (e.g. code, multimedia, and/or collaborative works). Even where you have permission, make sure you adhere to any conditions (say, for licenses). Provide references, links, and comments to include attribution related to your sources wherever possible - its social research etiquette!

In particular, the following are some of the things you can do to meet these criteria:

### Group supported work

You might:

* Start by giving more context to the task – for example, create a section to frame the problem: e.g., explaining what the program does and how it might be used.
* Add appropriate names, headings, and/or subheadings as well as markdown cells and comments to reflect your understanding of the machine learning workflow.
  + Explain what each cell in the code is for. If you know what each line does, then great. But you are not expected to know this.
  + Better to demonstrate a general understanding of the role of each cell in relation to the high level machine learning pipeline:
    - Getting the data.
    - Exploring the data.
    - Preparing the data for ML.
    - Selecting and training the model.
    - Evaluation and interpretation of the results.
  + The notebooks already have high level headings to facilitate your task.
* You are allowed to:
  + Confer with class groups.
  + Use AI tools (e.g. HuggingChat).
  + Employ general search engines.
  + Reference Keras webpages.
  + Look on stack overflow.

To help you understand the code in more depth, but make sure you give credit where it is due.

* You are encouraged to collaborate. You are not expected to demonstrate your knowledge, but your ability to synthesise knowledge from different sources in a way that is accessible to the broader audience.

### Your Own Creative and Critical Content

* Exercise your creativity and critical thinking to write your own narrative to make it accessible and engaging for the broader arts and humanities audience. This also is supported by talking to others in your group and beyond.
* Feel free to embed images etc – let your imagination run free.
* Finally, include your own critical reflection on:
  + How you would evaluate the result of your generative AI code?
  + How you might apply your code to your own data :
    - What kind of data would you be interested in applying it to?
    - Why are you interested in doing that?
    - How do you think you would need to transform your content for the computer?
  + What ethical concerns you envision in relation to this code or a similar development.
  + **No code necessary for any of the above.**

## Designated Notebooks

In the [Lab Resources](https://moodle.gla.ac.uk/course/view.php?id=39566#section-3) section, you will find two Jupyter Notebooks linked:

* [**“(Option 1) Generating Chorales With Neural Networks”**](https://moodle.gla.ac.uk/mod/resource/view.php?id=4095657)
* [“**(Option 2)** **Generating Text With Neural Networks”**](https://moodle.gla.ac.uk/mod/resource/view.php?id=4095660)

## Workflow You Should Follow:

* Download one of the designated notebooks for Part 3. Do not work on both. Work on only one. However, a discussion with someone who has worked on the other could add interesting comparisons.
* Open Anaconda, launch Jupyter Notebook, and open the downloaded notebook and follow the instructions therein.
* Before you finish your lab, upload your notebook to your GitHub repository – even if you have not finished it.
* It is good practice to create multiple commit points, for example, after each **Task** or **Part** is completed with an appropriate message to record progress.
* Do not wait until all the lab notebooks and exercises are all ready and complete. This is not good information management practice and misses out on some of the advantages of using GitHub.

## Sources of Inspiration

To get inspired! Take a look at the following links:

* Google Doodle: [Celebrating Johann Sebastian Bach](https://www.google.com/doodles/celebrating-johann-sebastian-bach) and [Behind the Doodle: Johann Sebastian Bach](https://www.youtube.com/watch?v=XBfYPp6KF2g)
* [MusicLM: Generating Music from Text (paper and dataset)](https://google-research.github.io/seanet/musiclm/examples/) (Agostinelli et al 2023)
* Alternative text generation code from Google Colaboratory: [Text Generation with RNN](https://colab.research.google.com/github/tensorflow/text/blob/master/docs/tutorials/text_generation.ipynb)
* [Get Shakespeare Version of Your Text](https://shakespeare-text-generator.netlify.app/)
* [Keras](https://keras.io/getting_started/) pages – with code examples and explanations of commands in Keras and Tensorflow

# Summary

Congratulations on making it to the end of the final lab sheet for this course.

In this sheet you have reviewed the machine learning workflow, explored some neural networks on tensorflow playground, and learned about AI and ethics as well as explainability on Kaggle. You also looked at the Embedding Projector to explore questions of bias.

Finally, by presenting the code in at least one notebook for generative AI, you deepened your understanding of the machine learning workflow.

# **Information on Your Assessment**

Your Portfolio forms 50% of your final grade.

This Portfolio consists of:

1. Your notebooks and content resulting from your lab tasks of Week 1-3.
2. Any notebooks resulting from this document for Weeks 5-11, where the tasks of Part 3 are carried out on your chosen notebook.
3. Following the processes in the Week 4 lab to complete the notebooks above.

Note:

* Pay attention to the guidance on what and how you submit it as detailed in the [Portfolio Guide](https://moodle.gla.ac.uk/mod/page/view.php?id=3851020). It is you responsibility to follow the instructions properly.
* In particular,
  + **Make sure you submit your GitHub URL** at the link AI for the [Arts & Humanities: Share Your GitHub Repository URL](https://forms.office.com/e/LSK58WQt8k).
  + **And submit the required documents on the Moodle as per the guide.**

I look forward to receiving your submissions in December!