**SEP 788/789 – Deep Learning and Neural Networks**

Project Description

Hamidreza Mahyar

[mahyarh@mcmaster.ca](mailto:mahyarh@mcmaster.ca)

Winter 2022

Introduction

As part of the SEP 788/789 course, students need to complete one hands-on project in a group of 2-3 students. Each group should pick their project from the list given below. Each deep learning project is adapted from a Kaggle competition or other well-known dataset, where a challenge (from a specific industry) for an open-source dataset is stated. Students need to form a group by 1st Feb 2022. Students without a group will be assigned to the group automatically by Avenue to learn at the end of the day on 1st Feb 2022.

# Project 1: Music generation

Source: [Music Dataset](https://github.com/umaniamir/music_dataset)

Description

Machine learning has covered almost all the industries where it has helped humans to mimic human-like behavior. Additionally, there are quite a few tasks where ML algorithms are proved better than humans where humans have physical limitations to achieve the task. The music industry is also not left behind. A few efforts are going on where AI and ML have improved the overall music experience by auto-generating music, correction of music notes, retrieval of information from music, etc. Here, the Music generation task is evaluated to generate beautiful music that has never been played before.

This dataset contains excellent, peaceful music from a piano that everyone likes. Here, the task is to generate beautiful music that has never been played before using deep learning approaches. The dataset contains two objects, notes, and chords. This project aims to train this dataset to create a model that predicts music from random notes and chords.

Dataset is split into two components: chords and notes. You need to use the music21 library to parse the mid file and get chords and notes. Notes contain pitch, Octave, and offset, and chords are containers for a set of notes played simultaneously.

More details regarding parsing data structure and organization can be found on the dataset [music21](http://web.mit.edu/music21/).

Project Outcomes

* Pre-process the data and create the dataset into trainable components.
* Evaluate various algorithms which can generate the music.
* Train a model to generate music from given random notes and chords.

# Project 2: Fake news detection

Source: [Fake news dataset](https://github.com/umaniamir/FakeNews_Dataset)

Description

Social media has provided an excellent interactive technology platform that allows the creation, sharing, and exchange of interests, ideas, or information via virtual networks very quickly. A new platform enables endless opportunities for marketing to reach new or existing customers. However, it has also opened the devil’s door for falsified information which has no proven source of information, facts, or quotes. It is really hard to detect whether the given news or information is correct or not. Here, as a part of this project, we need to detect the authenticity of given news using DL.

The dataset contains around 7k fake news, including a title, body, and label (FAKE or REAL). The task is to train the data to predict if the given news is fake or real.

Project Outcomes

* Pre-process the data to remove stop words. Stop words are the most occurring words in the language. It’s necessary to filter that out first.
* Evaluate the various algorithms which can affect the best outcome
* Train a model to predict the likelihood of REAL news.

# Project 3: Visual Object Tracking

Source: [Visual Tracking Dataset](http://cvlab.hanyang.ac.kr/tracker_benchmark/datasets.html)

Description

The topic of self-driving cars has been at the forefront of many legislative debates and is being included in many newer car models; Tesla, in particular, has a reputation for many car models with reliable self-driving features. One such focused area is visual tracking of the objects. Visual tracking is the ability to focus on one object that is stationary or having a movement.  The visual tracking technique is derived from humans as it mimics how a child learns. This ability progresses when a child visually follows people or objects and later learns to move eyes instead of the head. This technique is helpful to derive the relation of an object concerning the environment, Object detection, coordination of different tasks, etc. Visual tracking is quite an enjoyable task considering the number of objects in the field of view (FOV) of the camera. This project is about tracking an object into the field of view using any DL methods, benchmarking the dataset, and identifying the best metrics that can efficiently track the object.

This dataset contains sequences of images from recent literature that have one or more objects in it. If more than one kind of object is available, the dataset should contain identifier+’.’ +id\_number. For example BlurCar.1 BlurCar.2.

See more information at the dataset [website](http://cvlab.hanyang.ac.kr/tracker_benchmark/datasets.html).

Project Outcomes

* Identify everyday objects such as cars and roads within the image
* Explore the various algorithms to track the object and find the best one.
* Train the model to track the object on the screen.
* Predict the tracking of the object using a training model.

# Project 4: Object detection for Self-driving Vehicles on adverse conditions

Source: [Object detection of Cars](http://cadcd.uwaterloo.ca/)

Description

The topic of self-driving cars has been at the forefront of many legislative debates and is being included in many newer car models; Tesla, in particular, has a reputation for many car models with reliable self-driving features. This being said, accurate image and object recognition from a live dash feed is no easy feat; when human lives and safety are involved, the proper operation of a vehicle must always be prioritized – be it by man or machine. Significant problems with these models are to cover all the possible scenarios which can occur during life. There are quite a few adverse situations that occur quite often, i.e., rain, snowfall, night images; in the dataset included in this project proposal, you will undertake the task of object detection from a car’s dash capture in various adverse situations.

This dataset provides data images of cars in various street views. The image perspective is taken from a driver’s/dashcam point of view. Each image with cars includes bounding boxes to outline the location of the vehicle in the picture. For an AI model to be used for self-driving purposes, its algorithm must achieve high accuracy and fast prediction for self-driving/real-time monitoring purposes.

Project Outcomes

* Identify/detect cars in a variety of views
* Test your trained model for accuracy and speed – can it be used on the live video feed and allow a machine to react to high-speed, real-time information?

# Project 5: ChatBot

Source: [Questions Answer Dataset](https://www.cs.cmu.edu/~ark/QA-data/?ref=hackernoon.com)

Description

Machine learning has covered almost the entire industry, where it has helped humans to mimic human-like behavior. Additionally, there are quite a few tasks where ML algorithms are proved better than humans. One such progressive area is Human Assistance. Human assistance is AI technology that can understand customer needs and provide relevant information. Many companies have started using Human assistance as Customer support systems that can reduce human efforts for satisfying customer needs. Here are a few big names in the market who use human assistance like Alexa developed by Amazon, Siri by Apple, Cortana by Microsoft, etc. A chatBot is a chat-based customer support system where AI can assist customers in providing information in their interested areas.

The dataset contains questions and answers related to Wikipedia articles. This project intends to train the Natural Language Processing (NLP) based model to provide best-suited answers from the given articles. More information about the dataset can be found on the [website](https://www.cs.cmu.edu/~ark/QA-data/?ref=hackernoon.com).

Project Outcomes

* Explore various ways that data can be trained
* Identify the best model to train the data
* Train the natural language processing-based model to predict the anonymous questions
* Predict the answers to new questions that are not part of training.

# Project 6: Lane curve detection

Source: [Curved Lane Dataset](https://xingangpan.github.io/projects/CULane.html)

Description

Autonomous driving is in boom during recent advancements where vehicles need to move safely with little or no human interaction. Quite a few efforts are going on where companies can generate advanced algorithms which can detect the objects into the driver’s live dash. For a self-driving car to function accurately, it needs to detect lanes and traffic lights related to the current vehicle and objects in the line of sight. Lane detection is one of the essential areas where precision matters the most. As a part of this project, lane detection is being evaluated, which can help make more accurate decisions during self-driving.

This dataset contains a large scale of images for traffic lane detection. It has more than 55 hours of videos and 133,235 frames. The dataset is divided into various categories like Normal, Crowded, Night, No Line, Shadow, Arrow, Dazzle light, Curve, Cross Road

For this project, you most definitely will need to read up more about the dataset on the [website](https://xingangpan.github.io/projects/CULane.html).

Project Outcomes

* Develop scenes from data and identify lanes
* Explore various computer vision and machine learning algorithms to detect lanes and identify the best met the hod.
* Use a trained model to identify lanes
* Use positional data of things within a scene to identify how off the vehicle is concerning the center off the line.

Project 7: Personality Prediction

Source: [Myers-Briggs Dataset](https://www.kaggle.com/datasnaek/mbti-type)

Description

With the growth of industry 4.0, the need for talent management is at its peak. Many companies are identifying ways to get the best talent in the market. It’s been said that the company that can acquire a good talent will win the race. Having said that talent management is not an easy task and there are many ways companies identify talent as good. One very popular mechanism is a personality test using Myers Briggs test. In personality typology, the Myers–Briggs Type Indicator (MBTI) is a self-reporting questionnaire indicating differing psychological preferences in how people perceive the world and make decisions. The test attempts to assign four categories: introversion or extraversion, sensing or intuition, thinking or feeling, judging or perceiving. One letter from each category is taken to produce a four-letter test result, such as "INTJ" or "ESFP". So, for example, someone who prefers introversion, intuition, thinking, and perceiving would be labeled an INTP in the MBTI system. Details can be found [here](https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/) and [here](https://en.wikipedia.org/wiki/Myers%E2%80%93Briggs_Type_Indicator).

This dataset contains over 8600 rows of data, on each row is a person’s:

* Type (This person 4 letter MBTI code/type)
* A section of each of the last 50 things they have posted (Each entry separated by "|||" (3 pipe characters))

For this project, the aim is to identify the personality type of a person from the given post on the internet.

Project Outcomes

* Explore the various ways for data preprocessing
* Explore various ways that data can be trained
* Identify the best model to train the data
* Train the natural language processing-based model to predict the personality type
* Predict the answers to new questions that are not part of training.

Project 8: Sign Language recognition

Source: [*ASL Alphabet Images*](https://www.kaggle.com/kuzivakwashe/significant-asl-sign-language-alphabet-dataset) *and* [*ASL Alphabet Videos*](https://aslbricks.org/New/ASL-Videos/)

Description

Sign languages are visual languages that use a hand, facial and body movements as a means of communication. There are over 135 different sign languages all around the world including American Sign Language (ASL), Australian Sign Language (Auslan), and British Sign Language (BSL).  Sign languages are an extremely important communication tool for many deaf and hard-of-hearing people.  It’s not just about hand movement but also body and facial expressions. Facial expressions in sign languages can express both emotion and grammatical information. For instance, eyebrows can be raised and lowered to change the structure of a sentence from a statement to a question. It’s been said that people listen with eyes with sign language.

Here, there are two datasets given as per below. You need to choose one approach.

1. ASL alphabet image (<https://www.kaggle.com/kuzivakwashe/significant-asl-sign-language-alphabet-dataset>)

In this approach, there are plenty of images provided per alphabet of the English language. The student has to predict the alphabet from the image.

1. A large-scale dataset for Word-Level American Sign Language(<https://aslbricks.org/New/ASL-Videos/>). This dataset contains plenty of videos related to the word. Each video is in word. The task is to train this video to predict the word.

Project Outcomes

* Explore the various ways for data preprocessing
* Explore various ways that data can be trained
* Identify the best model to train the data
* Train the model to predict the alphabet or a word based on which dataset you choose.
* Predict the new alphabet or a word that is not part of training.

Project 9: Fake currency detection

Source: *[Currency dataset](https://archive.ics.uci.edu/ml/datasets/banknote+authentication)*

Description

It is well known that we live in an age where financial and technology sectors are growing hand in hand. This joint growth is often attributed to the fintech industry which is leveraging traditional banking services with technologies like artificial intelligence, blockchain, mobile services, etc. Great technological advancement leads to new ways of fraud. One such concerning area is Fake currency. Recently, advancements in the printing and scanning industry made the counterfeiting problem grow more vigorously. As a result, counterfeit currency affects the economy and reduces the value of original money. Thus, it is most needed to detect fake currency.

Data were extracted from images that were taken from genuine and forged banknote-like specimens. For digitization, an industrial camera usually used for print inspection was used. The final images have 400x 400 pixels. Due to the object lens and distance to the investigated object grayscale pictures with a resolution of about 660 dpi were gained. Wavelet Transform tool was used to extract features from images.

Attribute Information:

1. variance of Wavelet Transformed image (continuous)  
2. skewness of Wavelet Transformed image (continuous)  
3. kurtosis of Wavelet Transformed image (continuous)  
4. entropy of image (continuous)  
5. class (integer)

Project Outcomes

* Explore the various ways for data preprocessing
* Explore various ways that data can be trained
* Identify the best model to train the data
* Train the model to predict the fake note

Project 10: Landmark detection

Source: [Landmark Dataset](https://www.kaggle.com/google/google-landmarks-dataset)

Description

The Internet has broken down communication barriers between cultures in a way that could only be dreamed of in earlier generations. The effect of that is the increase in global tourism over the past decade. That leads to mass visiting famous places or landmarks in the world which might be difficult without globalization. Landmark Detection is the task of detecting popular man-made sculptures, structures, and monuments within an image. This technology can predict landmark labels directly from image pixels, to help people better understand and organize their photo collections. Today, a great obstacle to landmark recognition research is the lack of large annotated datasets. This motivated us to release Google-Landmarks, the largest worldwide dataset to date, to foster progress in this problem.

The dataset contains URLs of images that are publicly available online (this [Python script](https://www.kaggle.com/tobwey/landmark-recognition-challenge-image-downloader) may be useful to download the images). Note that no image data is released, only URLs.

The dataset contains test images, training images, and index images. The test images are used in both tasks: for the recognition task, a landmark label may be predicted for each test image; for the retrieval task, relevant index images may be retrieved for each test image. The training images are associated with landmark labels and can be used to train models for the recognition and retrieval challenges (for a visualization of the geographic distribution of training images, see [3]). The index images are used in the retrieval task, composing the set from which images should be retrieved.

Project Outcomes

* Explore the various ways for data preprocessing
* Explore various ways that data can be trained
* Identify the best model to train the data
* Train the model to detect the landmark

# Deliverables

Each group should deliver:

1. Reports (15%)
   1. Report 1: problem statement understanding (Feb 20th, 2022)

* Gather domain information regarding the problem statement from Google.
* Ask so many questions, what & why is this problem?
* See/visualize the data – Histogram, box and whisker’s plot, Scatter plot,  Pair plot, etc.
* Given the label column, is this a classification or regression task? And why? Explain it in detail. (Students should add the theory behind this in detail.)
* Plot the distribution of data (using sklearn/Pandas).
* play with the data and choose your ML approach (given your understanding of the problem in report 1)
  1. Report 2: Data pre-processing and choice of approach to implement (Mar 6th, 2022)
* Data Pre-processing:
  + - * + Normalize the data,
        + Handle missing values,
        + Check for outliers,
        + Balance vs. imbalanced dataset,
        + Features selection
* Evaluate different ML models in detail (mentioning their pros/cons).
  1. Report 3: Explanation of the method(s) and competing approaches (Apr 2nd, 2022)
* Choose one approach with proper reason.
* Fit your model, fine-tune hyper-params, plot learning curves for train, Val, and test data, report the results in test data, etc.
* Mention your conclusion.
* What are the limitations of the model?
* What are the future steps to improve the problem statement?

Note:

* The final report should look like a minimum of 5-6 pages (double column on each page). That means each report should have around 2 pages of technical content where students need to add as much information as possible. i.e. Diagrams, equations, tables, etc. Please let us know if you need more information.
* Please submit only new content for each report submission mentioned above except the report 3. The report 3 or a final report should contain a complete report with all the content of the previous reports.

1. Code (10%) (Apr 9, 2022)
   1. Codes must be clean, commented, easy-to-read, executable
   2. Create a GitHub repo and commit your codes there (share the link to the repo in reports)
   3. Codes will be reviewed using an AI-powered tool, so be careful plagiarism
   4. The deadline for your last commit is Apr 9th, 2022
2. Class Presentation (15%) (Apr 9, 2022)
   1. Problem statement
   2. Your proposed approach
   3. Results
   4. Presentation

# Pick your project

Each group (2-3 students) should send us an email [mahyarh@mcmaster.ca, CC: umania@mcmaster.ca] (subject line starts with “SEP 788/789”) and rank the projects in the order you would like to work on them (e.g., a project in rank #1 is the project you would like to work on the most and the project in rank #5 means you are least interested in). Then, based on your preference and popularity distribution of projects, we will assign a project to your group.