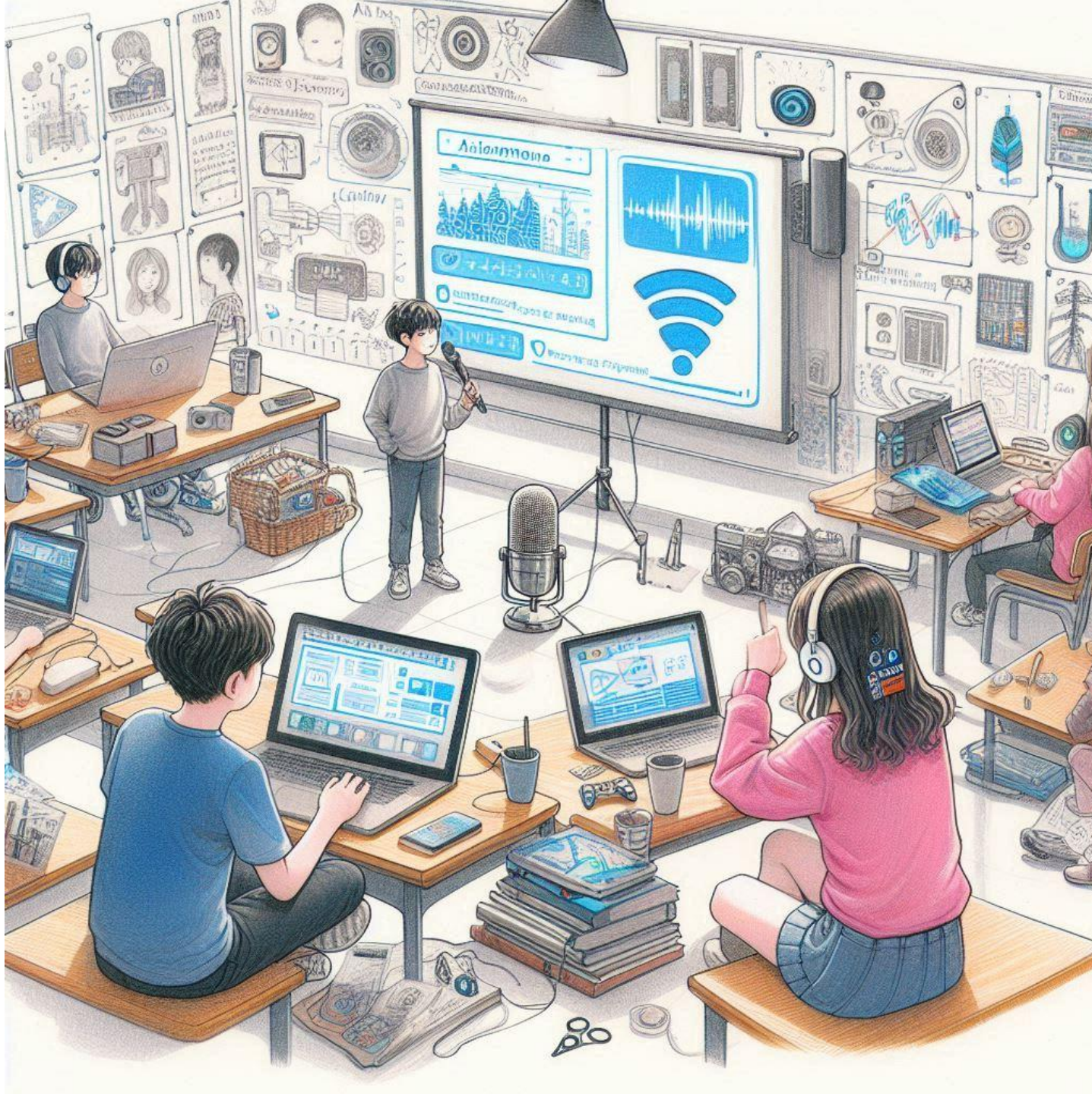
Signal and Natural Language Processing with Deep Learning

Assignments

Course 2024



Proposed Assignments

1. Music Genre Classification
2. Classical Music Genre Classification
3. EDM Genre Classification
4. Sentiment Analysis in written English (or your mother tonge)
5. Sentiment Analysis in spoken English (or your mother tonge)
6. DeepFake Voice Detection
7. Speaker identification/recognition (Biometric authentication system)
8. Speaker Diarization
9. Easy English-Spanish translator (written language)
10. Text Summarization using Transformers
11. Sound classification using Transformers
12. SPAM SMS classification
13. Fine Tunning an open source LLM for a specific topic (to define the topic)
14. Open topic (to be accepted)

Guidelines

1. Select 3 topic preferences using Google Form
2. Star looking for information on Internet
   - State of the art
   - Algorithms
   - Examples
   - Database for training
3. Write code for training a DNN
4. Write a report
5. Upload report & code to PoliformaT / Resources
6. Present a PowerPoint in class
   - Explaining the development
   - Presenting demos, if possible
   - About 25 minutes
   - Scheduled for December

Links & Resources

- Below are some links, references and databases that you can start from
- They are just the beginning to help you not start in blank, but you should look further
- You can always ask your teacher beforehand with any questions, queries or suggestions

General

- Hugging Face Datasets
- https://huggingface.co/docs/hub/datasets
- Tensorflow Datasets
- https://www.tensorflow.org/datasets/catalog/overview?hl=es-419
- Kaggle Datasers
- https://www.kaggle.com/datasets

- Hugging Face Model Hub
- https://huggingface.co/docs/hub/models
- Audio models
- https://huggingface.co/docs/transformers/model_doc/audio-spectrogram-transformer
- Kaggle Models
- https://www.kaggle.com/models

1. Music Genre Classification

- Find databases of songs by genres
- Expand the database: Rip CD's, Spotify, P2P
- https://en.wikipedia.org/wiki/List_of_classical_music_genres
- Genres:
  - Pop
  - Rock
  - Blues
  - Heavy Metal
  - Jazz
  - Latin (tango, bossa nova, ballroom, etc)
  - Salsa / Merengue / Bachata / Ballenato
  - Reguetón
  - Classical

2. Classical Music Genre Classification

- Find databases of songs by genres
- Expand the database: Rip CD's, Spotify, P2P
- https://en.wikipedia.org/wiki/List_of_classical_music_genres
- Genres:
  - Medieval
  - Renaissance
  - Baroque
  - Classical
  - Romantic
  - 20th Century
  - Film score …

3. EDM Genre Classification

- Find databases of songs by genres
- Expand the database: Rip CD's, Spotify, P2P
- https://www.armadamusic.com/news/edm-electronic-dance-music
- Genres:
  - Big Room
  - Chill-Out
  - Deep House
  - House
  - Progressive House
  - Electro house
  - Hardcore
  - Drum and bass
  - Techno
  - Trance
  - Vocal Trance

4. Sentiment Analysis (written)

- Sentiment Analysis inspects the given text and identifies the prevailing emotional opinion within the text, especially to determine a writer's attitude as positive, negative, or neutral.
- https://github.com/bonigarcia/nlp-examples/blob/master/4.Neural_NLP/2_sentiment_analysis.ipynb

5. Sentiment Analysis (spoken)

- Speech sentiment analysis is a technology that gauges the emotional tone of spoken language by analyzing tone of voice, word choice, and other linguistic cues.
- Intro
- https://medium.com/@alyzehkazmi/sentiment-analysis-from-audio-an-exploration-of-different-machine-learning-algorithms-6ea1359d08d6

- Emo-db Database (in German)
- http://www.emodb.bilderbar.info/download/
- Code
- https://github.com/Vijayvj1/Sentiment-Analysis-On-Voice-Data
- Microsoft API
- https://github.com/mrako/speech-sentiment-analysis
- List of Datasets and Models
- https://github.com/nehith23/Speech-Sentiment-Analysis

6. DeepFake Voice Detection

- An exhaustive list of resources (very good)
- https://github.com/media-sec-lab/Audio-Deepfake-Detection

- Dataset
- https://www.kaggle.com/datasets/birdy654/deep-voice-deepfake-voice-recognition

- Code
- https://github.com/sksmta/audio-deepfake-detection
- https://github.com/Srujan-rai/Deepfake_voice_detection

- Web APP
- https://github.com/Jerald-Golden/Audio-Deepfake-Detection

7. Speaker recognition/identification (Biometric authentication system)

- Speaker Recognition is speaker verification and identification software that distinguishes individuals using their unique voice characteristics.

- VOXCELEB (1251 speakers, 150 000 samples)
- https://www.tensorflow.org/datasets/catalog/voxceleb?hl=es-419

- Code
- https://github.com/donnydazzler/group-voice-biometrics?tab=readme-ov-file
- https://github.com/rvsolanki97/Speaker_identification_Biometric_using_Voiceit2_API_2.0

8. Speaker Diarization

- Speaker diarization is the process of segmenting audio recordings by speaker labels and aims to answer the question “who spoke when?”.

- Overview
- https://lajavaness.medium.com/speaker-diarization-an-introductory-overview-c070a3bfea70

- Code
- https://huggingface.co/pyannote/speaker-diarization
- https://github.com/taylorlu/Speaker-Diarization

- VOXCELEB (1251 speakers, 150 000 samples)
- https://www.tensorflow.org/datasets/catalog/voxceleb?hl=es-419

9. Easy English-Spanish translator (written language)

- Easy-Translate
- Easy-Translate is designed to be as easy as possible for **beginners** and as **seamless** and **customizable** as possible for advanced users
- https://github.com/ikergarcia1996/Easy-Translate

- Other Code
- https://github.com/devjwsong/transformer-translator-pytorch

10. Text Summarization using Transformers

- Intro
- https://towardsdatascience.com/text-summarization-using-deep-neural-networks-e7ee7521d804
- https://github.com/Tian312/awesome-text-summarization
- https://github.com/Storiesbyharshit/Natural-Language-Processing/tree/master/Text-Summarization-using-Transformers-T5

11. Sound classification using Transformers

- BEATs
- https://arxiv.org/pdf/2212.09058

- AST: Audio Spectrogram Transformer
- https://paperswithcode.com/paper/ast-audio-spectrogram-transformer
- ss

12. SPAM SMS classification

- SMS Spam Collection Dataset
- https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset

- Code: BEGGINER SMS-Spam-NLP
- https://www.kaggle.com/code/selimincekara/begginer-sms-spam-nlp-nb-recall-90
- Code
- https://www.kaggle.com/code/aleyhere/sms-spam-classifier-beginner-guide
- Code using Transformers
- https://www.geeksforgeeks.org/sms-spam-detection-using-tensorflow-in-python/
- Model
- https://huggingface.co/Sanrove/albert-spam-sms-classification-finetuned

13. Fine Tunning an open source LLM for a specific topic

- Topic to be defined in collaboration with the teacher
- https://labelyourdata.com/articles/llm-fine-tuning/top-llm-tools-for-fine-tuning
- https://www.youtube.com/watch?v=eC6Hd1hFvos
- https://www.youtube.com/watch?v=gs-IDg-FoIQ
- https://www.youtube.com/watch?v=iOdFUJiB0Zc
- https://www.youtube.com/watch?v=mrKuDK9dGlg

14. Open topic (to be accepted)

- You can propose a topic not listed above, motivating your interest in it and suggesting a source of information.
- It must be accepted by the teacher.