**Ten possible problems that could be address in the course**

**Idea 1: Music Recommender System**

Creating a social network for musician and music enthusiasts and ensuring that music content on the platform is of interest to users is a key to ensuring the sustainability of the music app. This idea is about developing a music recommender system, which can constantly recommend similar, yet refreshing music to the preference of the user. This concept revolves around advanced audio processing techniques, specifically converting music into vector space and conducting similarity analysis. Unlike conventional recommender models, our approach focuses on integrating the analysis of audio content without user-based factors. An additional experiment involves applying Fast Fourier Transform (FFT) to the audio files, as part of the data pre-processing. Similarity check of songs can be performed with FAISS. Spotify data can be used for a recommender system, then running the data through an existing model should give us similarities between song. For proof of concept, it is maybe realistic to aim for the accurate classification of song genres based on the audio only. Ethical considerations would include the need to find publicly permissioned data and audio song tracks. If user-based factors are included, there would be concerns surrounding data privacy.

**Idea 2: Single Feature Extraction from Music**

Train a model that gets a song or a track as input, and output a melody with single instrument, just like human would listen to a song (with drums, guitars, saxophone, and many more instrument at the same time) and whistle a single melody corresponding to this song. This idea has potential application in music search (e.g. enhanced Shazam) and can be used to complement idea 1. Producing data can be tricky. A method to create the dataset could be to pair songs with their cover versions, to attempt to map two or more different songs into the same vector. The melody of both the cover versions and actual song can be extracted for similarity comparison and training of GenAI model.

**Idea 3: Touch Screen Surface Interaction**

Utilize a webcam to enable any surface to function as a touch screen. Develop a model that enables the user to view a augmented reality input device on the table surface, detects the user taps on the surface via a camera and converts it to input data for an application. This can be utilized for various purposes, such as a piano keyboard, number pad, or drawing pad. An advanced version could enlarge the phone view for visually impaired users projecting the phone's content onto a desk, wall, or any flat surface and interacting with it as if it were a touch screen. This project can make use of OpenCV and MediaPipe for the hand detection and tracking. The data required for training of models can be produced by recording our fingers while typing. Each letter captured is the label of the observation.

**Idea 4: Sign Language Speech to Text**

Developing a model to convert signed speech to text could bridge the gap in communication between a deaf and a hearing person (with no knowledge of sign language). Currently, there have been [models](https://huggingface.co/RavenOnur/Sign-Language) capable of [classifying images of signed alphabets to its corresponding letter](https://www.sciencedirect.com/science/article/pii/S2667305323001096). Research work done previously have suggested the use of different image and video processing techniques to increase the accuracy of the model. This would be an interesting problem to tackle to try to increase the accuracy of existing models for longer signed sentences. There are some existing datasets which can be tapped on for the project: [American Sign Language Lexicon Video Dataset (ASLLVD)](https://www.bu.edu/asllrp/av/dai-asllvd.html), [A large-scale dataset for Word-Level American Sign Language](https://github.com/dxli94/WLASL), [Sign Language MINST](https://www.kaggle.com/datasets/datamunge/sign-language-mnist/data?select=sign_mnist_train) etc. Data can also be extracted from videos containing sign language captions using existing audio and video models. However, this might be challenging, especially considering the lag between spoken words and the sign language interpreter. An ethical consideration to take note of would be that accuracy is important in translation to avoid miscommunication and misunderstanding between users. If video data collected during the translation process is use in training the data further, this could lead to concerns surrounding privacy and personal data.

**Idea 5: Cryptocurrency and Stock Market Prediction Model with Expert Insights**

Develop a machine learning model to address the multibillion-dollar question: how to invest wisely? Instead of relying solely on complex algorithms, combine expert opinions with data analysis. Utilize predictions from financial analysts who frequently write articles or publish videos on YouTube. The algorithm assigns a current geopolitical and economic situation to a few analysts who have a proven track record in that situation, for example, analyst who accurately predicted the impact of a crisis in the Chinese real estate market on American stocks will be selected during times of a similar crisis. By doing so, the model provides a list of articles or videos that are likely worth checking out, allowing users to make more informed investment decisions by themselves.

**Idea 6: Fish Swimming Movement Analysis**

Tracking the behaviour, patterns and anomalies in fish swimming movement, crossing that with weather data may detect fish reaction to natural hazards. The data for this project consists of about 100 fishes in Þingvallavatn with tags that send location every 10 to 30 seconds (with a lot of missing data when the receiver is out of range).

**Idea 7: Horse / Dog Breed Classifier**

Even though model to classify Horse or Dog breeds have been developed, the challenge lies in enhancing the performance of such ML models or experimenting with different models (eg. RCNN, Yolo) to get a higher accuracy rate. The data required for this classification task for [horses](https://www.kaggle.com/datasets/olgabelitskaya/horse-breeds) or [dogs](https://www.kaggle.com/datasets/jessicali9530/stanford-dogs-dataset) can be found online.

**Idea 8: Dairy Farming Assistance**

Currently, costly necklaces with sensors are being used to track cows in the cowshed, monitor each cow's eating habits, steps taken per day, and whether she is in heat. If this can be achieved using simple cameras instead, it could be more cost-effective and scalable. This idea would require the use of image and video recognition models (eg. Lucas Kanade Point Tracking Algorithm, Background filtering). However, data for this project might be difficult to obtain as for the scope of this semester.

**Idea 9: Classification of Fake and Genuine Reviews**

With the rising presence of bots being used to shape public opinion about products, it is becoming increasingly difficult to differentiate genuine review from fake ones. Developing a model with the use of NLP can help users make more informed decision about their purchase. The model can be trained on this [dataset](https://www.kaggle.com/datasets/mexwell/fake-reviews-dataset).

**Idea 10:** **Automated Recipe Generation from a Picture**

Creating a recipe generator app which allows users to simply snap a picture of their desired dish to obtain the recipe for it. To ensure the dish is tailored to the user’s preference, users have the option to indicate the current ingredients they have at hand, preferences, dietary restriction, cuisine, etc. This project will combine the use of Image Recognition models (eg. Yolo) to detect the type of dish and Natural Language Processing to return a step-by-step recipe guide for users.

**Disclaimer:** ChatGPT 3.5 was used for the checking the existence of certain ideas and the cosmetic improvements to English writing in the report.