Coursework Description

Application of AI methods for multimodal heart disease diagnostic

Background

Health applications are one of the fields in which artificial intelligence will play a major role. Not only can monitoring patients' health be studied, but also preventive actions aimed at the early detection of diseases. Therefore, different types of data, including records of age, weight, blood type and pressure, diet type, smoking, etc., can be used to detect the risk of developing illnesses, such as cardiovascular diseases.

In addition, image processing applied to radiography can be used to detect certain anomalies, such as heart anomalies. Therefore, both modalities, tabular data and X-ray images, can be processed with artificial intelligence techniques to indicate the presence or the potential risk of developing such diseases. However, this tool cannot replace medical knowledge and expertise and must be limited to a decision-support role. Furthermore, the analysis of input data using AI techniques can provide a better understanding of the role of each characteristic in the development of the diagnosed disease(s).

Scope and tasks

Using two datasets, one for tabular data of patient health about cardiovascular diseases, and the other for X-ray images about heart anomalies, we aim to develop artificial intelligence techniques to detect the presence of each disease from the corresponding datasets.

Task1- heart disease severity classification using tabular data

Using **Dataset1**, concerning patient health records, the data are presented in tabular form. The aim is to use the data present as features and labels, to train a model capable of detecting the presence or the level of the illness (from 0: no illness to 4: severe illness). Therefore, our first aim is to build:

a) A binary-classification model, to detect the presence of a heart disease. In this case, any label in the last column (*level*) above 0 is considered as illness.

b) A multi-class classification model, to detect the level of the heart illness, using a multi-class classifier.

However, not all features (column 2:age) to (column 15: thal) could be equally useful. Therefore, **feature selection** can be taken into consideration during the development of the models. Besides, **feature normalisation or standardisation**, may be required to reduce the data sparsity.

Input features

The following list contains a description of the different features and their values:

- **id:** (Unique id for each patient)
- **age:** (Age of the patient in years)
- **sex:** (Male/Female)
- **origin:** (place of study)
- **cp:** chest pain type: 1. typical angina; 2. atypical angina; 3. non-anginal; 4. asymptomatic
- **trestbps:** resting blood pressure (resting blood pressure (in mm Hg on admission to the hospital))
- **chol:** (serum cholesterol in mg/dl)
- **fbs:** (if fasting blood sugar > 120 mg/dl)
- **restecg:** (resting electrocardiographic results)
- **Values:** [normal, stt abnormality, lv hypertrophy]
- **thalach:** maximum heart rate achieved
- **exang:** exercise-induced angina (True/ False)
- **oldpeak:** ST depression induced by exercise relative to rest
- **slope:** the slope of the peak exercise ST segment
- **ca:** number of major vessels (0-3) colored by fluoroscopy
- **thal:** [normal; fixed defect; reversible defect]

Note! The first column (Id) is not considered as an input feature.

Output labels

The output label is:

- The last column "level" indicates the predicted attribute (0 shows no disease and 1, 2, 3 and 4 shows different level of disease)
- This label can used as it is for multiclass classification. However, for binary classification, labels must be converted into two values (0/false: if the original label is 0, and 1/True: if the original label is above 0).

Required AI techniques

To carry this project, it is required to follow these guidelines:

- Use more than one classification model, compare their performances for each classification task (binary or multiclass).
- Classification models should firstly be implemented with the methods studied in this module, namely Decision trees, naïve Bayesian, Random forests, linear classifiers, linear SVM, Artificial Neural Networks and Deep Neural Networks.
- Other classification methods can be used (optionally), however only for comparison with the studied methods.
- To improve the performance, you can use different architectures or settings for each method.
- Performance must be assessed using standard evaluation metrics, such global accuracy, and class-wise precision, recall and f1-score.
- Findings must be discussed and commented.

Task 2- Heart disease detection using X-ray images

Using **Dataset 2**, containing X-ray images of the chest, we aim to classify them into presenting signs of heart disease or not.

The dataset comprises 17 folders of images, each containing images of the related disease, i.e. {*Abscess, Ards, Atelectasis, Atherosclerosis of the aorta, Cardiomegaly, Emphysema, Fracture, Hydropneumothorax, Hydrothorax, Pneumonia, Pneumosclerosis, Post inflammatory changes, Post traumatic ribs deformation, Sarcoidosis, Scoliosis, Tuberculosis and Venous congestion*}.

However, only the following are directly related to heart anomalies:

- ✓ Atherosclerosis of the aorta
- ✓ Cardiomegaly
- ✓ Pneumonia
- ✓ Pneumosclerosis
- ✓ Post-inflammatory changes
- ✓ Post-traumatic ribs deformity

Therefore, we aim to implement **a binary classifier** to classify images into **heart-related** or **non-heart-related diseases**, as explained above.

However, the input must be the **raw image**, without any prior feature extraction, and therefore the most suitable AI technique must be selected, e.g. **convolutional neural networks (CNN)**.

Task 3- Image reconstruction and denoising

Also using **Dataset2** for X-ray images, it would be interesting to exploit generative models, such as Simple or Variational Autoencoders (VAE) and/or Generative-Adversarial Networks (GAN) to improve the quality of images by reconstruction from clean (original) images or from noisy images (by adding an artificial noise). This can provide a model that helps to use low-quality images by applying reconstruction/enhancement before classification.

Problem statement

Diagnosis of disease using multimodal data, i.e. medical records and/or radiographic images, can be approached as a classification problem:

- Collect the datasets (see the AI_CW2_2025-2026_Material folder in Canvas/Learning Material), and analyse the data contained in each one (tabular medical records in the Dataset1 and radiographic images in Dataset2). The data need to be divided into training and test sets. When necessary, carry out the necessary preprocessing steps, such as normalisation, standardisation or cleaning.
- For Dataset1, about patient health data, develop several classification models as mentioned above in **Task1.**
- For the X-ray image dataset, develop:
    - A classification model in which the input is the raw image and the output is the heart-related / non-heart-related disease label, (see the list of heart-related diseases in the description of **Task2**)

    - A generative model, in which the input is a low-quality or a noisy image and the output is a clean image. In this case, noise can be artificially added to the images in the dataset, as a random Gaussian noise (see description of **Task3**).
- The evaluation of developed models must be included as follows:
    - For classification models (developed on each dataset), the confusion matrix and underlying evaluation metrics such as overall accuracy and precision per class, recall and F1 score. An interpretation of each measure must be provided.
    - For the image reconstruction/enhancement model, assess the quality of denoising by measuring the corresponding metrics, including the absolute and the relative root mean square error.
- Implement a program that allows the user to enter the patient's tabular data and/or uploading an X-ray image, as input, to provides the detected disease(s) as output, using the developed models. The program may also apply a reconstruction step to clean the upload images, before performing classification.
- Discuss the results obtained, making recommendations for further investigation and better implementation.
- Show how the developed models can have an impact on accelerating the health monitoring process and the precautions to be taken to avoid misuse of patient data or any other ethical issues related to the application of AI in the healthcare sector.
- Write a report of **1000 (at minimum) to 1500 (at maximum) words** to present your work and the results obtained
- Present your code as a **python notebook** (*.ipynb file*).

Note! Do Not include the datasets in the submitted work

keywords: Healthcare, heart disease diagnosis, patient data, X-ray images, AI, binary classification, multi-class classification, generative models, evaluation metrics, ethical issues, privacy.

Useful Reading

- T. Zhu, K. Li, P. Herrero and P. Georgiou, "Deep Learning for Diabetes: A Systematic Review," in *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2744-2757, July 2021, doi: 10.1109/JBHI.2020.3040225.
- Ahsan, Md Manjurul, and Zahed Siddique. "Machine learning-based heart disease diagnosis: A systematic literature review." *Artificial Intelligence in Medicine* 128 (2022): 102289.
- Rahman, Tawsifur, et al. "Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray." *Applied Sciences* 10.9 (2020): 3233.

Report structure and formatting

Report structure

You report of about 1000 (minimum) to 1500 (maximum) words should contain the information listed below.

- **Cover Page (1 page)**: Title, author, affiliation of the author, date, and abstract

- **Body (3 to 5 pages)**
    - **Introduction**: A rough overview of the disease diagnosis application you are tackling and its importance, frequently formulated to attract the reader's interest to the report.
    - **Background**: Here you are required to have an overview about the relevant methods.
    - **Methodology and Data**: This section contains an explanation, demonstration, description of the architecture of the models to be developed, or some interesting implementation techniques. Discussion of some targeted methods for solving the AI-based disease diagnosis problem will be mostly encouraged. Description of the knowledge representation (datasets used and data processing issues) are discussed in this part. The development of the AI system is discussed in this part as well
    - **Analysis and Discussions**: You need to present your findings and discuss, analyse, evaluate and/or criticise what you implemented and described in the previous part. You can also show the impact of such an application on accelerating patient health monitoring. Besides, the ethical issues resulting from the misuse of AI in patient disease diagnosis need to be reported.
    - **Conclusions and suggestions for future work:** the major findings from doing this coursework will be presented in this section. You can comment on the lessons learnt from doing this coursework, advantages/limitations of your AI solution, resources needed to implement the full system, what would you differently if you had more time, etc.
- **Bibliography and Citations (1 page ONLY)**: It is imperative that whenever you refer to a fact of some sort, you cite an authoritative source for that fact; most frequently, these sources will be scientific articles.

Report formatting

The report must be written in Word or any other Office tool, and saved as .doc, .docx or .pdf file. It must be formatted as follows:

- **Font:** Times New Roman
- **Font size:**
    - **Heading1:** 14 pts
    - **Heading2 and below:** 12 pts
    - **Heading3 and below:** 12 pts
    - **Body: 12pts**
- **Spacing:** Single
- **Layout:**
    - **Orientation:** Portrait**,**
    - **Columns:** One-column
    - **Margins:** Normal **(**2.54 cm from each side)
    - **Size:** A4 (21 cm x 29.7 cm)

Plagiarism and use of AI

- **Plagiarism policy**
    1. Submissions with a similarity (plagiarism) rate over 20% will also receive a mark of *zero*.
    2. To check the plagiarism rate, the report must be submitted separately (not included in the zip file containing the code) as a .doc or .pdf file

3. While submitting your report on Canvas, if you obtain a plagiarism rate above 20%, you can amend your report and submit again, as multiple submission will be allowed until the deadline mentioned above.

- **AI-Generated content**
  1. Using AI to generate code, text or figures is not allowed.
  2. The exam board will check, and if necessary, question the student(s) if there are any suspicion about AI-generated content included.

Requirements

Effort required

You are expected to demonstrate good effort in exploring the background of this problem, understanding the datasets, processing the data, developing meaningful AI models and implementing some basic aspects, and writing up your report. The coursework is worth 70% of the marks for this 20-credit module. A 20-credit module is nominally 200 hours of your time, some of which is spent in lectures, labs and private study. You should expect to be spending around 70 hours on this coursework.

Deadline for submission of report

The report must be submitted through Canvas before **6.00 pm Friday 16th Jan 2026**
Late submissions (without acceptable extenuating circumstances) will receive a mark of *zero*.

Specific Learning outcomes

This coursework will enable you to apply basic principles of AI in solutions that require problem solving, inference, planning, and learning. It will enable you to demonstrate an ability to share in discussions of AI, its current scope and limitations, and societal implications.

Marking Scheme

Criteria	>80%	70-80%	60-70%	50-60%	40-50%	39-25%	<25%
Presentation of the report (25%)	Fully adheres to student guidelines. Complete lack of spelling and grammatical errors. Excellent use of appropriate and scientific language. Excellent structure and organisation. Excellent readability.	Good degree of adherence to student guidelines. Lack of spelling and grammatical errors. Good use of appropriate language. Excellent structure and organisation. Good readability	Mainly adheres to student guidelines. Minor spelling and grammatical errors. Good use of appropriate language. Well-structured with logical organisation.	Some adherences to student guidelines. Some spelling and grammatical errors. Inconsistent use of appropriate language. Organisation and progression evident.	Little adherence to student guidelines. Many spelling and grammatical errors. Minimal use of appropriate language. Inadequate attention to structure and organisation	Does not adhere to student guidelines. Major deficiencies in spelling and grammar. Lack of appropriate language. A disorganised report with lack of evident structure.	No report submitted. Paperwork submitted does not constitute a meaningful report.
Description and achievement (25%)	Novel or innovative AI-based solution. Exceptional amount of high-quality work, which could include comprehensive testing, or critical reviews,	The AI system presented is meaningful, innovative challenging and possibly complex. Problem explicitly stated with precise explanation of	A relevant and original topic which is effectively translated into project aims and objectives which are clearly stated. Good analysis and	Appropriate problem area chosen. Objectives outlines with the main areas of investigation identified. Some analysis and recommendatio ns.	Limited topic choice with the problem area poorly defined. Objectives vague or insufficient. Very limited analysis and recommendatio ns.	Simple or unoriginal problem showing lack of imagination. No analysis and recommendatio ns	No software/project completed. No working software/project (does not apply to theoretical submission).

	or results incorporated into prototype.	all research objectives. Excellent analysis and recommendations.	recommendations.		.		
The demonstration of development capability (25%)	Signs of professionally developed AI concepts, excellent understanding of user requirements, excellent specifications and system requirements. Specification effectively interpreted with signs of original thinking and/or research.	Fully developed AI concepts, fully developed user requirement , fully developed specification and System requirement. Evidence of excellent understanding of AI-development process and applications.	Evidence of developed AI concepts, developed user requirements, developed specifications and System requirements. Good understanding of the AI-development process and applications.	Partial in-depth development of AI concepts, Partial in-depth developed user requirements, Partial in-depth developed specifications and System requirements. Evidence of a fair understanding of the AI-development process and applications.	Limited development of AI concepts, limited developed user requirements, limited developed specifications and System requirements. Evidence of a limited understanding of the AI-development process and applications.	No or very limited development of AI concepts, no or very limited developed user requirements, no or very limited developed specifications and System requirements. No evidence of understanding of the AI-development process and applications.	No demonstration given. System failed to work. Student could not explain working of software or technical aspects of project (does not apply to theoretical submission).
Quality of work and understanding (25%)	Wide range of material investigated. Excellent evidence of critical evaluation and	Good range of material. Good evidence of critical evaluation. Good number of appropriate	Focus on key areas using relevant sources. Good range of references with a varied bibliography.	Adequate information survey with some evidence of investigation of a key area. Appropriate	Limited sources of information used. Limited range of references. Limited bibliography.	Review of existing literature not evident. No references provided with limited or	No understanding demonstrated. No conclusions. No future work suggested.

	original thinking. Substantial number of appropriate references. Evaluation and recommendations fully and appropriately reviewed and presented. Excellent link to implementation. Research findings fully consider broader issues.	references. Recommendations fully and appropriately reviewed and presented. Good discussion of implementation. Research findings discusses broader issues.	Clear recommendations identifying key issues. Research findings apparent with some consideration of broader issues.	range of references. Good bibliography. Recommendations identify some key issues. Research findings lack consideration of broader issues.	Limited recommendations. Lacking clarification of research findings.	omitted bibliography. No recommendations. Conclusions do not link to research findings.	