David Normoyle

K00141953

The benefits and drawbacks of automated information retrieval in Healthcare. How data mining can be used for diagnosis and to predict future health.

1. Summary [to be completed later]

This will be the last section, it is to be completed when the project is finished.

1. Introduction

For this topic it was important to look at what technologies are currently being developed in the Medical Technology sector. In an ever-expanding market it is important to look at what Is involved with bringing new technologies forward and the difficulties with regulations for newer and emerging technologies. This paper will look at Artificial Intelligence and Deep Learning Programmes, to test the viability of machine aided diagnosis/artificial intelligence as a diagnostic aid through data mining software.

Artificial intelligence defined in the Collins dictionary

“Artificial intelligence is a type of computer [technology](https://www.collinsdictionary.com/dictionary/english/technology) which is [concerned](https://www.collinsdictionary.com/dictionary/english/concern) with making [machines](https://www.collinsdictionary.com/dictionary/english/machine) work in an [intelligent](https://www.collinsdictionary.com/dictionary/english/intelligent) way, similar to the way that the human [mind](https://www.collinsdictionary.com/dictionary/english/mind) works. The [abbreviation](https://www.collinsdictionary.com/dictionary/english/abbreviation) [AI](https://www.collinsdictionary.com/dictionary/english/ai) is [also](https://www.collinsdictionary.com/dictionary/english/also) used.”

The aim is to write a code using computer software for coding languages Python/Spyder.

To create databases with questions and different variables that a user can interact with.

The programme could be used to look for patterns in family history of disease to try and predict possible health complications in later life. Considering variables such as age, family history, smoker/non-smoker, dietary choices (poor, balanced and extreme/unusual diet like vegetarian and vegan, unusual meaning outside a normal balanced diet), activity levels, alcohol use.

A script that can act as a questionnaire a doctor might use in hospital. For example, a person suffering from chronic infections can have their data analysed to check for a pattern to see what may be triggering infection. A person with a heart condition can have their data checked against other people who have the same condition, again to look for more patterns and similarities. To identify risk factors based from a users input.

Machine learning can be used to help interoperate this kind of data at a much faster level. If done correctly it has the potential to be very useful for gathering and understanding data and information.

Data definition according to the Collins dictionary, in terms of AI; “**Data** is information that can be [stored](https://www.collinsdictionary.com/dictionary/english/store) and used by a computer program.”

Information in terms of AI; “Information consists of the facts and [figures](https://www.collinsdictionary.com/dictionary/english/figure_1) that are [stored](https://www.collinsdictionary.com/dictionary/english/store) and used by a computer [program](https://www.collinsdictionary.com/dictionary/english/program).”

A programme that could do the same based on health information and history would be very beneficial in terms of saving time when compiling research into disease causes and treatments. To be able to retrieve more data from Electronic health Records.

1. Needs and Problems

I want to write a script that will have a login feature and be able to generate and store information that a user inputs.

A series of questions will be asked that the user can answer and this information will then be stored into a database.

When the questionnaire is finished the information will be stored and be ready for use again later to be interoperated.

A secondary script will them be used to look for correlating factors and to identify the highest risk factors as inputted by the user such as diet alcohol use smoking etc etc.

A tool that can help with the idea of evidence-based medicine. If a system can look through and interrupt data quickly and at a large rate it would be beneficial to studies when looking for correlation.

An example of this would be to look at diet factors and see how likely a person is to develop diabetes.

As more and more hospitals are likely to incorporate some kind of electronic health records then a system that can interrogate this data would be beneficial.

1. Goals & Objectives

I want to write a script that will have a login feature and be able to generate and store information that a user input.

A series of questions will be asked that the user can answer, this information will then be stored into a database.

When the questionnaire is finished the information will be stored and be ready for use again later to be interoperated.

A secondary script will them be used to look for correlating factors and to identify the highest risk factors as inputted by the user such as diet alcohol use smoking etc etc.

A tool that can help with the idea of evidence-based medicine. If a system can look through and interrupt data quickly and at a large rate it would be beneficial to studies when looking for correlation.

An example of this would be to look at diet factors and see how likely a person is to develop diabetes.

To create the script and have a database that it can look through.

1. Scope of work

To use python to create the script and to use Microsoft excel to generate data. As a proof of concept idea. I have decided to use Microsoft Excel to generate data as after many failed attempts to make python communicate with Microsoft Access have been unsuccessful. (Microsoft Access is a 32-bit programme Python is a 64 bit programme) and attempts at installing the 32-bit version of python have not worked.

The scope for this project is to see how data is handled and how it can be processed. To create a working script that for now will examine data in Microsoft Excel with the hope that in future it will be transferable to other types of database.

1. Time Management



1. Evaluation [To be evaluated and rechecked as it is written and from the study of more literature on this topic(Machine Learning and EHR]