**220COM: Coursework task 3 data mining:**

Data mining is the process of searching through a data set to discover patterns and behaviours within the data. It uses machine learning and intelligent methods to create models that we can then use. The aim of data mining is to analyse data that would usually not be analysed and only result in extra cost to store of no purpose. It can predict new behaviours and discover unknown patterns with a large set of data. Data mining is the most efficient way of making the most out data, “With the volume of data available today, organizations turn to Big Data management solutions and customer experience management solutions capable of advances data mining for translating raw data into actionable insights”(Molly Galetto).

An appropriate data mining application for a hospital is to detect is a patient has a high chance of developing cancer.

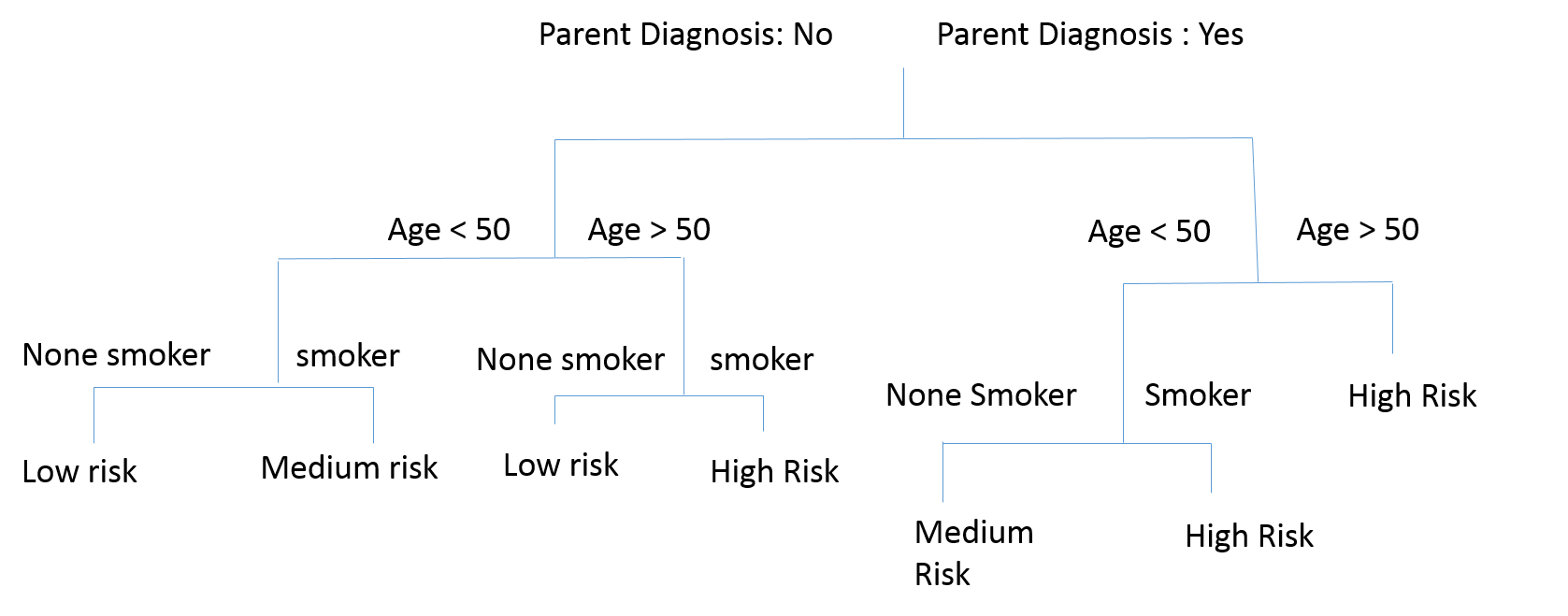
Problem definition: Using a patient’s data we need to detect is they have a high chance/risk of developing cancer.

Data Gathering and Preparation: Collect the patient data

e.g. Example set of data for one patient, if the system was implemented this would be completed and multiple patient data would be included.

|  |
| --- |
| Age: 17 |
| Parent diagnosis : Yes |
| Smoker: no |
|  |

Model Building and Evaluation: Decision Tree:



Use Knowledge: We now have an earlier warning system to assess whether a patient has a high, medium or low risk of cancer. Using the results we can consider how to improve the system. For example we could include more appropriate data for determining certain types of cancer, and use more patient medical details.

**References:**

GALETTO, M.

What is Data Mining? Learn about Definition and Purpose

“With the volume of data available today, organizations turn to Big data management solutions and customer experience management solutions capable of advanced data mining for translating raw data into actionable insights”

In-text: (Galetto 2016)

Galetto, M. (2016) What Is Data Mining? Learn About Definition And Purpose [online] available at:

<https://www.ngdata.com/what-is-data-mining/>