Responsible AI has gained significant attention due to the need for ethical, fair, and trustworthy AI systems. It addresses the potential risks and challenges associated with AI, promoting responsible and beneficial use of this technology. Explainability plays a crucial role in responsible AI. It helps to ensure that AI systems are transparent, accountable, and fair. Here are a few ways explainability helps in responsible AI.

One of the biggest reasons of explainability gaps are because the model selection process and feature selection process needs a strong understanding of the domain knowledge and understanding of complex business models. The only way the gap is addressed today is by sharing the model results and its top explaining variables. These variables are managed by data engineers and then used by data scientists to create other statistical features. Till the point the results reach to the business users, the features are even difficult to interpret by the business.

Responsible AI <- Model Explainability <- Model Output <- Model feature Creation & Creation <- Data variables

* Data variables are named in non-intuitive way. Confusing table names or even variables names are common problem in data management
* Data catalogues are designs for data systems and link management, lacks domain explainability
* Data scientists build statistical features and use technology driven methods to select models.
* Iterations between Data scientists and Business users have very little impact as even some of the explainability techniques are very naïve and uses the same feature names and columns names to explain model output.

One of the strongest usecase of LLMs for Data scientists, bridging the gap of domain, functional understanding while developing or experimenting with different models.

What can be a better dataset than a healthcare data to understand how it can simplify the explanation for a data scientist and in-turn improve explainability of the models for higher adoption.

I used an open-source data to understand if GenAI /LLMs can bridge the gap between data collection, data management, model selection, model explainability processes to improve business adoption.

I was really astonished to see how impactful it was!

I went from an output which looked like this,

Patient ID: ABC -> Model prediction: Malignant (80% probability) -> top 5 variables: [radius\_worst, area\_worst, compactness\_se, concavity\_worst, concavity\_se]

To this,

Patience ID ABC ->

For Patient ABC, the higher chances(80%) of malignancy of cancer is primarily caused by factors like larger variations and complexities in the length and boundary of the tumor, more severe concave portions, highest standard deviations of gray-scale values, larger sizes, and greater distances from the center to points on the perimeter of the tumor. These factors suggest a higher likelihood of aggressive and harmful characteristics in the cancer.

(Note, this is not validated by any professional expert, yet it certainly shows that a strong prompt engineering can make a solution more explainable, usable, adoptable, and hence responsible)

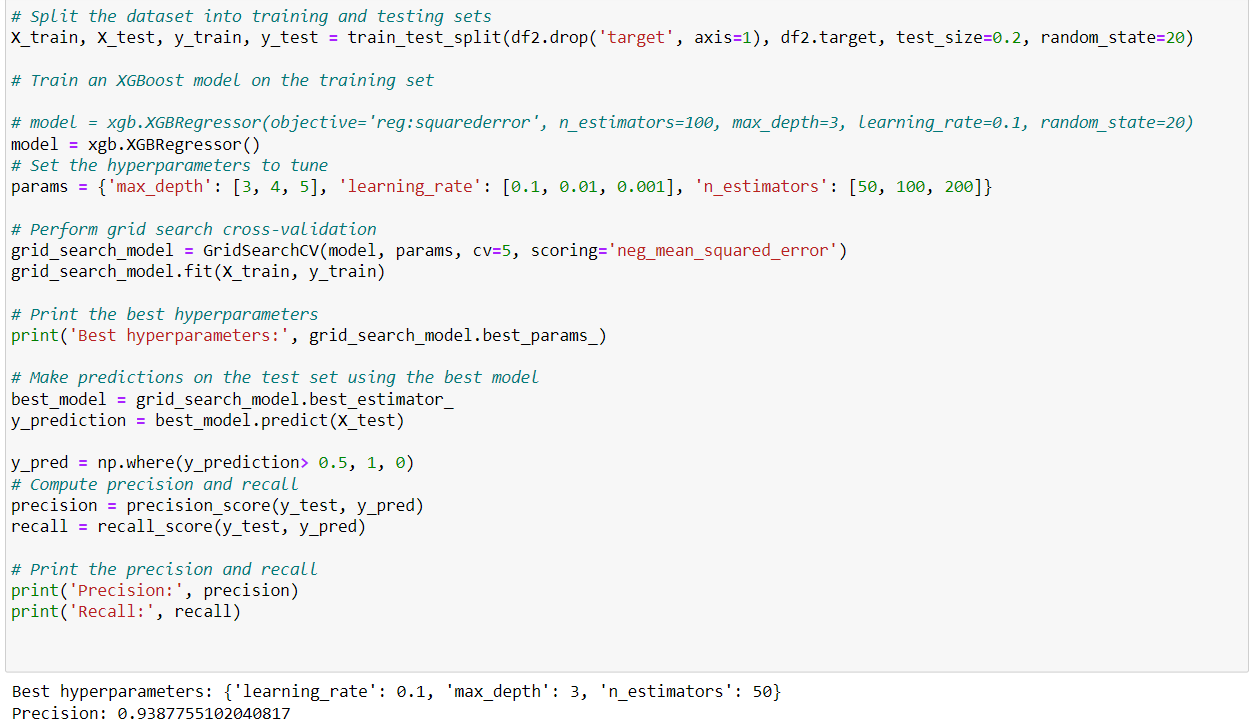
Here is how the power of CHATGPT was used.

Step 1: Data Understanding  
Data set is taken from [Breast Cancer Dataset | Kaggle](https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset). There are 32 variables and the first task is to understand the variables of the dataset. The objective is to build a classification model to uild classification models to predict whether the cancer type is Malignant or Benign. I am not focusing on the hyper parameter tuning part as the objective is to improve explainability with the existing model. Here is how the variable list looks like,

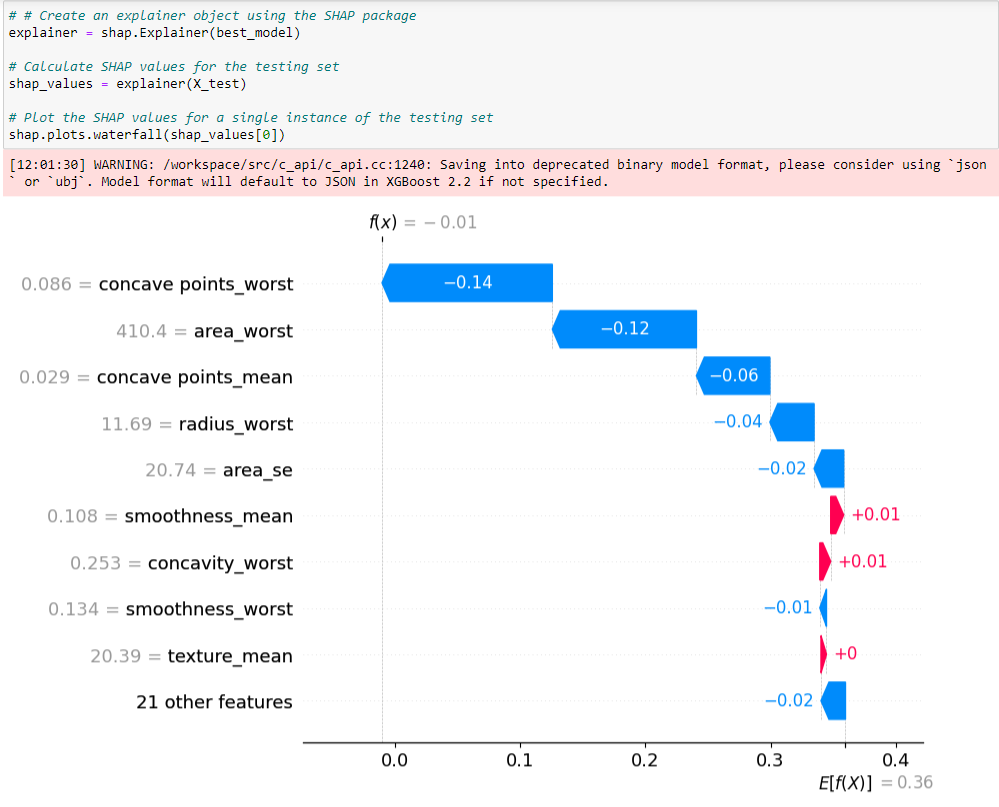


The variable names are very non intuitive. I went through multiple searches and website to understand meaning of each of the variables, but I wasn’t greatly successful. I did understand that there is a process of nuclear feature extraction for breast cancer diagnosis. Again, a scientific paper and I was more interested in building a model and see the accuracy of the model in predicting Malignant diagnosis.

Step 2; Model building and Model selection: I used cross validation of XGB model to find the best parameter and train the model. The accuracy of the model looks very good.



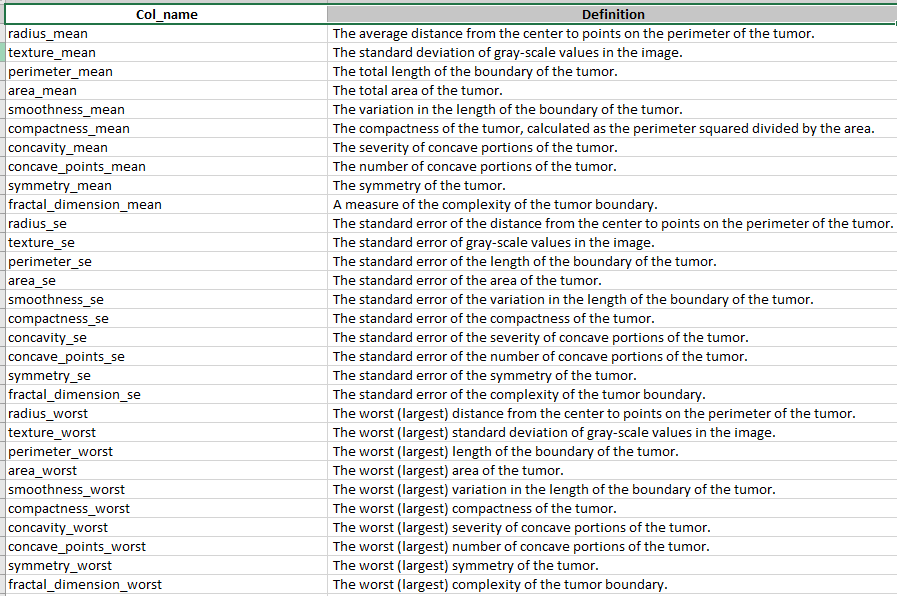
Step 3: Using Shap to get top 5 variables and Wondering if these variables make sense!!!



This is where I started understanding more about the data and how it is collected. I went through <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/1905/1/Nuclear-feature-extraction-for-breast-tumor-diagnosis/10.1117/12.148698.short?SSO=1> which explains how the data is collected. It had multiple medical terms which were not easy for me to understand.

Step 4: Using Chatgpt to understand as well as explain the variables.

4.1 – I provided the abstract of the paper to Chatgpt for it to learn and then asked it to Elaborate the column names and make it more readable/intuitive. ChatGPT did an amazing job of it. I was totally surprised.



4.2 – Use the Shap values, the definitions and a prompt to get an answer like a doctor.

Prompt#1: Output

**Summary Statement of High Severity Causes of Cancer**

As a specialist doctor, I have analyzed the provided table to provide a summary statement of the causes of cancer. The 'definition' column represents the cause, while the 'shap\_value' column showcases the severity.

Based on the data, the following high severity causes of cancer have been identified:

The worst (largest) severity of concave portions of the tumor - High severity

The worst (largest) area of the tumor - High severity

The worst (largest) distance from the center to points on the perimeter of the tumor - High severity

PromptEngineering #xx (After a few attempts)

**Summary: Causes of Malignant Cancer**

The malignancy of the cancer is primarily caused by factors such as larger variations and complexities in the length and boundary of the tumor, as well as more severe concave portions and higher standard deviations of gray-scale values. Additionally, larger sizes and distances from the center to points on the perimeter of the tumor contribute to the malignancy. These factors indicate a higher likelihood of aggressive and harmful characteristics in the cancer.

A fun example 😊

Summary: Conversation between doctor and patience explaining cause of cancer being malignant

Doctor:

Good afternoon, Mr. ABC. I have reviewed your medical reports and I would like to discuss the cause of your cancer with you. It's important to understand that cancer can have multiple causes, and in your case, we have identified a few factors that may have contributed to its development.

Mr. ABC:

Thank you, Doctor. I've been really worried about what might have caused my cancer. Can you please explain it to me?

Doctor:

Certainly, Mr. Smith. Based on the analysis of your medical data, the following factors have been identified as potential causes of your cancer:

The worst (largest) severity of concave portions of the tumor

The worst (largest) area of the tumor

The worst (largest) distance from the center to points on the perimeter of the tumor

It's important to note that these factors are identified based on statistical analysis and may not represent the sole cause of your cancer. Cancer is a complex disease and can have various contributing factors. It's always recommended to consult with your healthcare team for a comprehensive understanding of your specific case.

Mr. ABC:

I see. So, these factors played a significant role in the development of my cancer?

Doctor:

Yes, Mr. Smith. Based on the data analysis, these factors have shown a high severity in relation to the development of your cancer. However, it's important to understand that cancer is a multifactorial disease, and other factors may also have contributed. It's crucial to continue with the recommended treatment plan and follow-up appointments to ensure the best possible outcome.

Mr. ABC:

Thank you, Doctor, for explaining this to me. I appreciate your guidance and will continue with the treatment as advised.

Doctor:

You're welcome, Mr. Smith. I'm here to support you throughout your treatment journey. If you have any further questions or concerns, please don't hesitate to reach out. We'll work together to fight this cancer and improve your health.