**User Guide for “TDAcancer\_detect” Function**

**Overview**

The **TDAcancer\_detect** function is designed for cancer detection and identification of the tissue of origin, specifically distinguishing stomach cancer and pancreatic cancer from other cancer types. This function analyzes biomarker concentration levels obtained from a patient’s blood test by utilizing Topological Data Analysis (TDA). It extracts statistical information about the topological features from the patient data and compares these with normal and cancer-specific training datasets. Based on this analysis, the function determines the presence of cancer and identifies the tissue of origin if the cancer is either stomach or pancreatic.

**Usage**

The function can be used in a default mode with biomarker concentrations as inputs but also allows for extensive customization through various parameters.

**Input and Parameters**

* **Load the dataset:** Load the patients’ data sheet ‘Normal and Cancer’ from the provided Excel Data File ‘Clinical cancer data.xlsx’.
* **File\_path:** Path to the file containing the Excel file ‘Clinical cancer data.xlsx’.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Serial no. | Parameter | Values | Parameter Types | Remark |
| 1 | Regression Difference Percentage | 1 | Default | >1 condition employed in the Cancer detection model |
| 2 | Distance %(1) | 4 | Default | >2 condition employed in the Cancer detection model |
| 3 | 1-Dim Hole Max Life Range | 0.025, 0.055 | Default | <0.025 or >0.055 condition employed in the Cancer detection model |
| 4 | Distance %(2) | 14 | Default | <14 condition employed in Stomach Cancer detection model |
| 5 | Distance %(3) | 10 | Default | >10 condition employed in Pancreas Cancer detection model |

**Example Usage**

Here is an example command to run the **TDAcancer\_detect** function:

The user need to enter the patient’s biomarker concentrations obtained from the blood test in the form of excel file as like example dataset “clinical cancer data.xlsx”. The excel sheet contains a row named as “Enter the Patient blood test data here=>”. Enter the biomarker concentrations to the respective biomarker columns in the row and save the file. Next, provide the file path to the function and run it. Except the biomarker concentrations, all the parameters used in the function are default parameters. In this example, we chose a Stomach cancer patient which patient ID is ‘INDI 162’. we enter the biomarker concentrations for this particular example to the respective biomarker columns in the row named as “Enter the Patient blood test data here=>” in the attached excel sheet “clinical cancer data.xlsx”. We then enter the file\_path. The following command will be fine:

*Results = TDAcancer\_detect(file\_path)*

**Output**

The above function will return a Persistence Diagram of Stomach Cancer, a Persistence Barcode of Stomach Cancer, a Persistence Diagram of Pancreas Cancer, a csv file (“Clinical cancer data\_StomachDet.csv”) containing Stomach cancer detection details, and finally a csv file(“Clinical cancer data\_PancreasDet.csv”) containing Pancreas Cancer detection details. If something wrong happens, the error information will be displayed.

**Error Handling**

If there is an issue with the data input or during any of the processing steps, the function will return an error message detailing what went wrong, ensuring that users can make necessary adjustments.