Early detection of Pancreatic Cancer using Machine Learning

Mohammed Shehbaz

Project Aim & Objectives

‘To make use of supervised learning algorithms, in the form of a classification model to identify early signs of pancreatic cancer’

Data Assimilation and Analysis

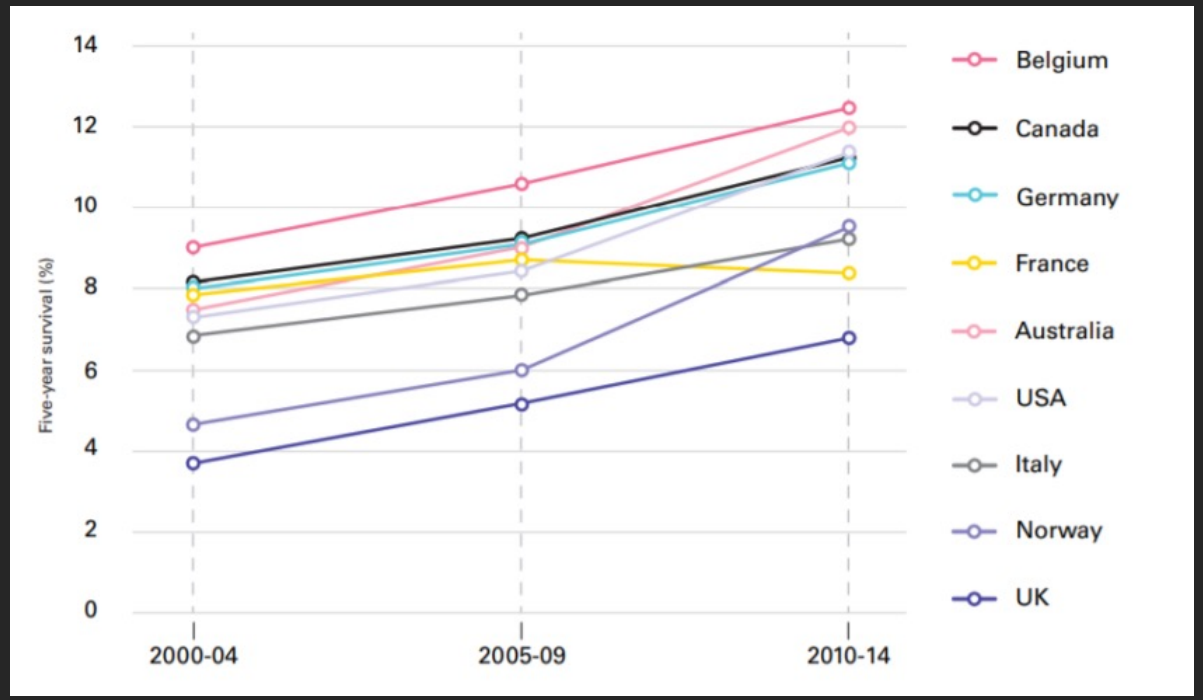
Data Pre-processing and Visualisation

Implementation of machine learning models

Evaluation of metrics and hyperparameter tuning

Project Importance

- Causes 9000 deaths in UK each year, 5th most deadly cancer
- Lowest survival rate of all common cancers, 5 year survival less than 7%. Five-year survival in Scotland across all persons is only 5.6% (2010-2014)
- **7 in 10** people with pancreatic cancer do not receive any active treatment
- 3 in 5 people are diagnosed at a late stage where potentially curative surgery is not possible



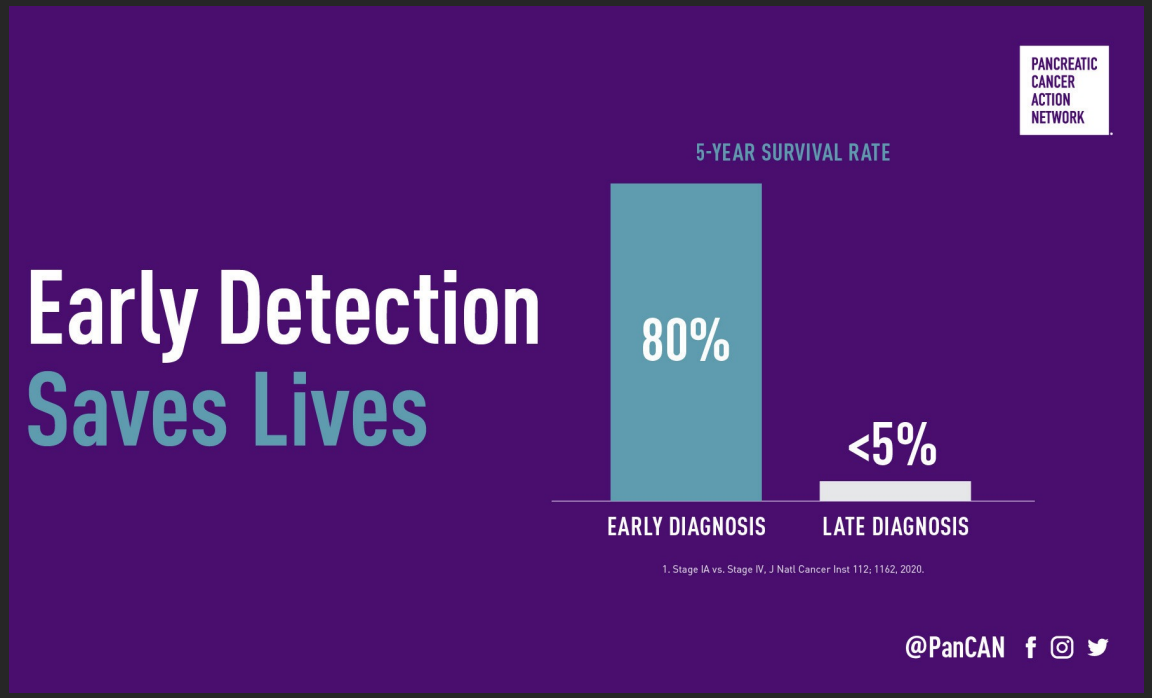
Current work in the field

Common diagnosis methods include

Image tests: create pictures of organs for visualisation (CT, MRI, PET)

Removing a tissue sample for testing: a sample of tissue is collected from the pancreas

Blood test: Doctor may test blood for specific proteins (tumour markers) shed by pancreatic cancer cells



Literature

Classification of premalignant pancreatic cancer mass-spectrometry data using decision tree ensembles

https://bmcbioinformatics.biomedcentral.com/track/pdf/10.1186/1471-2105-9-275.pdf

Machine learning of clinical performance in a pancreatic cancer database

https://www.sciencedirect.com/science/article/pii/S0933365710000527?casa_token=aHCls7ifAAAAAAAA:27IQnydBDCHfmLrhSv8StUzx49foGKIGNHog8M4KUYM0pxCOTRozQBPZywb4ryZ7zx1xEgfm

Early detection of pancreatic cancer

https://www.sciencedirect.com/science/article/abs/pii/S2468125319304169

Using probe electrospray ionization mass spectrometry and machine learning for detecting pancreatic cancer with high performance

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7013221/

A machine learning approach identified a diagnostic model for pancreatic cancer through using circulating microRNA signatures

https://reader.elsevier.com/reader/sd/pii/S1424390320306153?token=A60182E96B82C4299E2B11D8017E098446D8AE1DD11F8162FEA350F29239AE568A83A628C3BF1C56C52810E25D9FEF30&originRegion=eu-west-1&originCreation=20211219143623

Related Datasets

Urinary biomarkers for pancreatic cancer

https://www.kaggle.com/johnjdavisiv/urinary-biomarkers-for-pancreatic-cancer

Clinical Data Prediction Model to Identify Patients With Early-Stage Pancreatic Cancer

Optum de identified IER data set

Thank you for Listening!