**“Pancreatic Cancer Detection Using Machine Learning, Deep Learning, TabPFN, and Explainable AI”**

**Sanad Alawi¹ and Rami Al-Ouran²**

¹Student, Al Hussein Technical University, Amman 11831, Jordan  
²Department of Data Science and AI, Al Hussein Technical University, Amman 11831, Jordan

Corresponding author: Rami Al-Ouran (rami.alouran@htu.edu.jo)

**Abstract:** Pancreatic cancer remains one of the deadliest cancers due to its late diagnosis and the absence of effective, non-invasive early detection and identification methods.. In this research paper, we present an advanced AI-based diagnostic approach for pancreatic cancer detection using a real-world urinary biomarker dataset. Our methodology combines traditional machine learning (ML), deep learning (DL), and the recently released TabPFN model – an innovative tabular foundational model designed for accurate predictions on small datasets. We explore eleven different imputation techniques to handle missing data, evaluate 26 classification models, and perform extensive hyperparameter tuning using (a) GridSearchCV and (b) RandomizedSearchCV. Among all tested models, TabPFN demonstrated the most promising performance in terms of accuracy, robustness, and generalization. To ensure transparency and trust in predictions, we incorporate explainable AI (XAI) techniques, allowing insights into feature importance and decision behavior. Additionally, we conduct and analyze a short survey to gather complementary insights related to pancreatic cancer awareness and early detection challenges. Finally, the best-performing TabPFN model is deployed in a user-friendly web interface application, enabling clinicians or researchers to input patient biomarker values and receive real-time cancer risk predictions. This study demonstrates and shows how combining modern AI techniques with foundation models and medical biomarker data can support early ,pancreatic cancer detection and potentially improve patient outcomes.

**Keywords:** Pancreatic cancer detection, Urinary biomarkers, AI in healthcare, Machine learning models, Deep learning models, TabPFN, Explainable AI (XAI), TaPFN data augmentation, TabPFN-generated feature embeddings

# **1. Introduction**

**1.1 Computer Science: The Big Umbrella of Artificial Intelligence**

AI, standing for and representing Artificial Intelligence, is a multidisciplinary field of science, engineering, technology, mathematics, and computing, dedicated to developing intelligent machines and computer programs capable of performing tasks that typically and usually require human cognition. These tasks include reasoning, learning, decision-making, problem-solving, perception, recognition, understanding, planning, automation, prediction, and optimization (Russell and Norvig, 2016). AI systems leverage advanced algorithms, computational models, and data-driven approaches to enhance efficiency, adaptability, and intelligence across various different domains, including healthcare (Esteva *et al.*, 2019), finance (Danielsson and Uthemann, 2024), robotics (Levine *et al.*, 2018), and natural language processing (NLP) (Vaswani *et al.*, 2017).

Figure 1 below illustrates how Artificial Intelligence (AI) is a multidisciplinary field that integrates techniques from many different various domains, including mathematics, computer science (CS), robotics, and data science – among many others.

A diagram of a science

AI-generated content may be incorrect.**Figure 1 – Artificial Intelligence adopts techniques from various different domains**

However, Artificial Intelligence (AI) is just one of many specialized fields within the vast domain of computer science (CS). While AI has gained significant attention due to its transformative capabilities, it is only one component of the broader computer science (CS) umbrella, which encompasses various other critical disciplines. These include cybersecurity, networking, databases, software development, cloud computing, robotics, and data science, among others (*Map of Computer Science*, 2017). Each of these domains plays a vital (and crucial) role in advancing technology, securing digital infrastructure, and optimizing computational processes. As illustrated in Figure 2 given below, computer science is a multidisciplinary field that integrates multiple disciplines, with AI being just one of its many integral subfields.

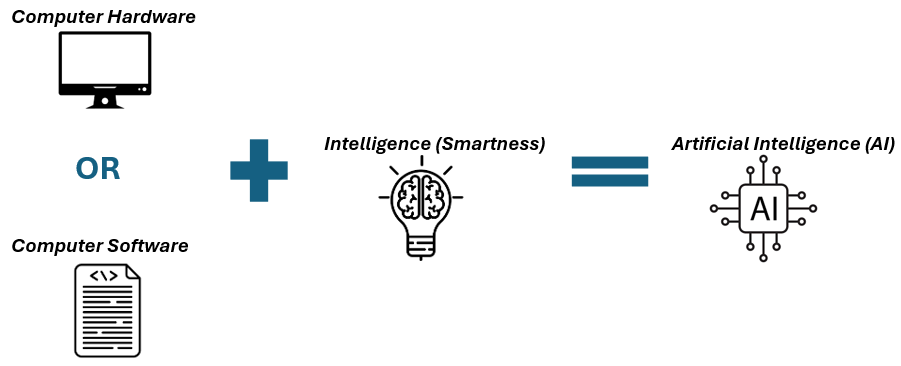
A chart of different colored squares

AI-generated content may be incorrect.**Figure 2 – The Broad Scope of Computer Science (CS) and Its Subfields**

**1.2 AI vs Traditional Programming**

Artificial Intelligence (AI) can be fundamentally understood as the combination of **computer hardware** or **computer software** with **intelligence** (smartness), resulting in a system capable of simulating human-like cognitive abilities (*What Is Artificial Intelligence? Definition, Uses, and Types*, 2024). As illustrated in **Figure 3** below, AI systems are created (and made) when computational resources (hardware or software) are enhanced with intelligent algorithms that enable and empower them to reason, learn, and make A diagram of a process

AI-generated content may be incorrect.decisions. This formula highlights (and shows) that AI is not just about programming machines to execute tasks – it’s about embedding cognitive abilities into systems, enabling them to analyze data, recognize patterns, and adapt to new – fresh information.

**Figure 3 – The Basic Formula Equation of Artificial Intelligence (AI)**

This concept represents a significant departure from traditional programming. Historically, programming computers involved writing a detailed and specific set of instructions for the machine to execute a particular – *specific* task. In this **traditional rule-based programming paradigm**, the programmer defines all possible scenarios and outcomes in advance. The system processes inputs according to these fixed rules and generates an output. If a new scenario arises, the program must be manually updated with additional rules – a time-consuming and *very* rigid process (*AI VS Traditional Programming - What’s the Difference?*, 2024).

**Figure 4** illustrates the fundamental differences between traditional programming and AI. In traditional programming, the process follows a structured flow where **inputs** and **rules** are explicitly defined, and the system produces an output based on these rules. If the input changes, the programmer needs to modify and change the rules manually to maintain (and keep) the desired outcome.

**Figure 4 – Comparison Between Traditional Programming and Artificial Intelligence (AI)**

In-contrast, AI-based systems reverse this approach. Instead of defining rules manually, AI models analyze data and identify patterns to generate the underlying rules. In an AI-based model, the system is provided with inputs and expected outputs during training. Through learning algorithms, the model identifies patterns and establishes its own decision-making framework. This allows AI systems to adapt dynamically to new data and improve their performance over time without any human intervention. As summarized in Figure 5 given below, AI-based systems outperform and surpass traditional programming in terms of adaptability, complexity, and scalability, enabling them to handle diverse and dynamic challenges more efficiently (and more effectively).

**Figure 5 – Key Differences in Processing Between Traditional and AI-Based Systems**

|  |  |  |
| --- | --- | --- |
| Aspect | Traditional Programming | AI-Based Systems |
| 1. Approach | Rule-based; relies on predefined instructions. | Learning-based; generates rules from given data. |
| 1. Adaptability | Fixed; requires manual updates to handle new cases. | Adaptive; improves automatically with new data. |
| 1. Complexity | Effective for well-defined tasks with limited variability. | Suitable for complex, high-dimensional, and dynamic tasks. |
| 1. Scalability | Poor; needs constant refinement for new scenarios. | High; improves with increased data exposure. |
| 1. Performance Over Time | Static; performance remains fixed unless manually improved. | Dynamic; performance improves as the model learns from new data. |
| 1. Example | A program for tax calculation based on fixed rules. | A model predicting heart disease risk based on patient health data. |

For example, a traditional program designed to identify animals in photos would require explicit rules for each animal type, such as “”””if the object has four legs and fur, classify it as a dog.”””” If a new breed or species is introduced, the program would fail unless the rules are manually updated and changed.

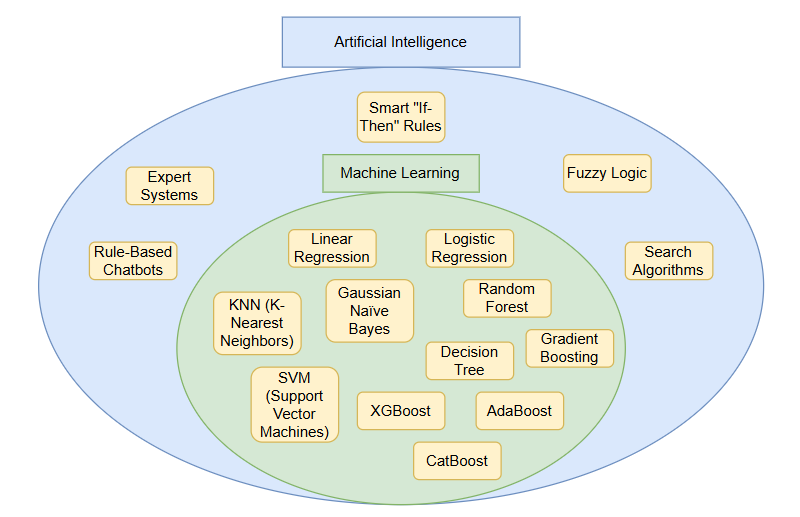
In contrast, an AI-based animal recognition system would be trained on thousands of images of animals. By identifying underlying patterns in the data, the AI model would learn to recognize animals based on shared features (e.g., fur texture, facial structure) without needing *any* explicit instructions. This adaptability makes AI particularly effective in fields like healthcare, where dynamic and complex patterns are common.

This shift from a rule-based to a learning-based paradigm underpins AI’s transformative impact on various industries. Traditional programming remains effective for well-defined tasks, but AI’s adaptability and learning capability make it more suitable for complex and evolving environments, such as: medical diagnosis, financial modeling, and natural language processing (NLP).

**1.3 AI vs Machine Learning**

Although the terms **Artificial Intelligence (AI)** and **Machine Learning (ML)** are often used interchangeably, they are not the same thing and refer to distinct concepts. AI is the broader field that encompasses any technique that enables machines to mimic human intelligence, including reasoning, problem-solving, learning, and decision-making (*Machine Learning vs. AI: Differences, Uses, and Benefits*, 2025). As shown in **Figure 6**, AI includes various methods such as: **Expert Systems**, **Smart “If-Then” Rules**, **Fuzzy Logic**, and **Search Algorithms** – which are not based on machine learning (ML) technology.

Machine Learning (ML), on the other hand, is a subset of AI that involves training models to learn patterns from data and improve over time without being explicitly programmed. ML methods rely on statistical techniques such as: **Linear Regression (LR)**, **Logistic Regression**, **K-Nearest Neighbors (KNN)**, **Support Vector Machines (SVM)**, **Decision Trees (DT)**, **Random Forest (RF)**, **Naïve Bayes (NB)**, and boosting methods like: **XGBoost (XGB)**, **CatBoost**, and **AdaBoost**. While all machine learning (ML) models are AI-based, not all AI techniques involve machine learning (ML).

**Figure 6 – Relationship Between Artificial Intelligence (AI) and Machine Learning (ML)**

**1.4 AI vs Machine Learning vs Deep Learning**

Similar to how **Machine Learning (ML)** is a subset of the broader field of **Artificial Intelligence (AI)**, **Deep Learning (DL)** is a specialized subset of ML. As shown in **Figure 7**, ML includes various methods for training models to learn patterns from data, including traditional statistical models such as: **Linear Regression (LR)**, **K-Nearest Neighbors (KNN)**, and **Support Vector Machines (SVM)** – which are not based on deep learning (DL) technology.

Deep Learning (DL), on the other hand, is a more specialized area within ML that involves training complex models inspired by the structure and function of the human brain. DL models rely on **neural networks** and are capable of learning hierarchical patterns from large datasets. Unlike most traditional ML models, DL methods rely more heavily on advanced training techniques such as: (a) **backpropagation** and (b) **gradient descent** to adjust and modify model parameters automatically during training (LeCun, Bengio and Hinton, 2015).

DL methods include various architectures such as: **Artificial Neural Networks (ANN)**, **Convolutional Neural Networks (CNN)** [used in computer vision], **Recurrent Neural Networks (RNN)** [used in sequential data], **Long Short-Term Memory (LSTM)** networks, **Transformers** [used in natural language processing], and **Generative Adversarial Networks (GANs)** [used for data generation]. While all deep learning (DL) models are machine learning-based, not all machine learning models involve deep learning (DL).

A diagram of a machine learning

AI-generated content may be incorrect.**Figure 7 – Relationship Between: Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL)**

**1.5 AI vs ML vs DL vs Explainable AI (XAI)**

Explainable Artificial Intelligence, also commonly referred to as **XAI**, is a rapidly emerging field within the broader scope of Artificial Intelligence (AI), dedicated to improving the **interpretability**, **transparency**, and **trustworthiness** of AI, Machine Learning (ML), and Deep Learning (DL) systems. As shown in **Figure 8** below, XAI intersects with AI, ML, and DL, providing insights into how complex models make predictions and decisions. Explainable AI includes a variety of methodologies and techniques – such as: **SHAP (Shapely Additive Explanations)**, **LIME (Local Interpretable Model-Agnostic Explanations)**, **Partial Dependency Plots (PDP)**, and **Feature Importance Analysis** – designed to enable human users to fully understand and trust the decisions produced by sophisticated – advanced AI *Artificial Intelligence* models (Adadi and Berrada, 2018).

Unlike traditional “”black-box”” models that generate accurate predictions without explaining how decisions are made, XAI aims to **illuminate (and brighten) the decision-making process** by offering clear explanations for model outputs. This interpretability is particularly important in sensitive and high-stakes fields such as **healthcare** (where AI influences diagnosis and treatment), **finance** (where AI shapes loan approvals and fraud detection), and **autonomous vehicles** (where safety and accountability are paramount). By enhancing transparency and accountability, XAI helps to build trust between AI systems and human users, ensuring (and making sure) that AI-driven decisions are *fully* aligned (and in agreement) with human expectations and ethical standards.

A diagram of a diagram

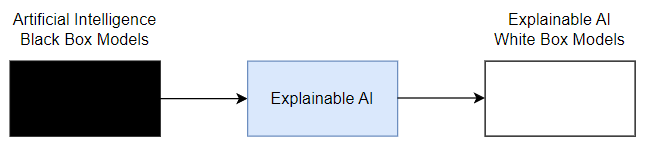
AI-generated content may be incorrect.**Figure 8 – Relationship Between AI, Machine Learning, Deep Learning, and Explainable AI**

**Figure 9** illustrates the increasing global interest in Explainable AI (XAI) over the past two decades, as reflected in Google Trends data. The x-axis represents the date (from 2004 to 2024), while the y-axis measures the level of interest based on search activity. From 2004 to 2016, interest in Explainable AI (XAI) remained minimal. However, starting in **2017**, there was a sharp increase in interest, likely driven by the rise of complex deep learning models and the growing need for interpretability. This upward trend reflects the expanding recognition of Explainable AI as a crucial component of AI development, especially in fields where transparency and accountability are essential. The consistent increase in interest through **2024** underscores the growing demand for interpretable AI models and highlights the critical (and crucial) role of XAI in building trust and improving the transparency of AI-driven decisions.

A graph showing a line

AI-generated content may be incorrect.**Figure 9 – Interest Over Time for “Explainable AI” Based on Google Trends**

**Figure 10** demonstrates the transition from traditional “black box” AI models to more transparent and interpretable “white box” models. Black box AI models often deliver high accuracy but lack transparency, making it difficult (and challenging) to understand how decisions are made. Explainable AI (XAI) serves as the bridge, introducing methods that provide insights into the decision-making process and enhance model interpretability. This shift allows AI models to transition into white box models, where the reasoning behind decisions is more accessible and understandable to human users, improving trust and accountability.

**Figure 10 – Transition from Artificial Intelligence (AI) Black Box Models to Explainable AI (XAI) White Box Models**

**1.6 ML vs DL vs Foundation Models (TabPFN)**

While both Machine Learning (ML) and Deep Learning (DL) have driven remarkable advancements in artificial intelligence (AI) over the past decade, their strengths differ significantly depending on the **type of data** and **problem domain**. Machine learning (ML) models, especially ensemble models such as: XGBoost (XGB), CatBoost, LightGBM, and Gradient Boosting (GB), have consistently outperformed deep learning (DL) models on **tabular datasets**, which are commonly found in real-world structured data applications. On the other hand, deep learning has achieved transformative results in **text** (e.g., ChatGPT, BERT), **image**, and **video processing** domains, thanks to its ability to learn hierarchical patterns through deep artificial neural networks (ANNs).

However, when it comes to **tabular data**, deep learning (DL) has traditionally underperformed compared to classical machine learning. This gap has led to the emergence of a new paradigm: **Foundation Models**. As shown in **Figure 11**, foundation models represent a major shift in AI development. Instead of being trained solely on a single dataset, they are pre-trained on **millions of synthetic tasks or massive corpora**, enabling them to generalize across many domains and perform well even with limited data. One of the most notable examples in the tabular domain is **TabPFN** – a cutting-edge tabular foundation model introduced by (Hollmann *et al.*, 2025).

**TabPFN** breaks conventional boundaries by leveraging a deep transformer-based *advanced* architecture trained on an enormous variety of simulated tabular problems. Once trained, it can make highly accurate predictions on new small-scale tabular datasets in a **zero-shot manner**, without requiring (or needing) any further tuning. This novel and innovative approach enables TabPFN to achieve levels of accuracy, robustness, and efficiency that were previously unattainable using standard – *typical* ML or DL methods.

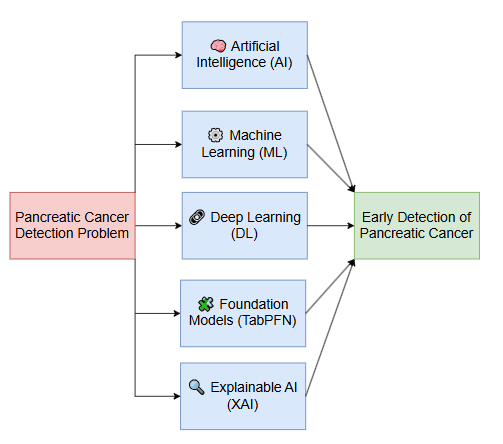
A diagram of a machine learning

AI-generated content may be incorrect.**Figure 11 – Relationship Between AI, ML, DL, XAI, and Foundation Models**

**1.7 Pancreatic Cancer Detection**

Pancreatic cancer is one of the most aggressive and lethal types of cancer, with a devastatingly low survival rate. According to (Debernardi *et al.*, 2020), only around **9% of patients survive beyond five years** after diagnosis. However, if detected at an early stage, the chances of successful treatment and long-term survival increase significantly. Unfortunately, most cases of pancreatic cancer remain asymptomatic until the disease has progressed to advanced, often incurable stages. This makes **early detection not just a clinical goal – but rather a life-saving necessity**.

Given the complexity, urgency, and subtlety of early-stage symptoms, traditional diagnostic methods often fall short. This is where Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Explainable AI (XAI), and Foundation Models such as 🡪 TabPFN can play a revolutionary role. By leveraging **non-invasive urinary biomarker data**, advanced AI techniques have the potential to uncover hidden patterns that may be missed (and ignored) by human experts. Given the critical importance of solving the pancreatic cancer detection problem, **Figure 12 illustrates the core objective of our research:** combining AI, ML, DL, TabPFN, and XAI techniques to achieve early and highly accurate detection of pancreatic cancer. Throughout this research paper, we aim to develop a solution that not only **maximizes diagnostic accuracy** but also ensures the process is **interpretable and transparent**, empowering doctors to understand and trust the model’s predictions – thus *thereby* improving patient outcomes and saving as many lives as possible.

**Figure 12 – AI-Based Framework for Addressing the Pancreatic Cancer Detection Challenge**

**1.8 Problem Definition**

The central problem addressed (and discussed) in this research paper is the early and accurate detection of pancreatic cancer – a disease known for its aggressive progression and extremely low survival rates. However, we do not aim to solve this problem using conventional or outdated methods. Instead, our goal and ambition is to critically review existing literature, identify current limitations, and propose a modern, innovative solution that leverages the latest cutting-edge advancements in Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Explainable AI (XAI), and Foundation Models such as: TabPFN.

Specifically, we aim to explore:

“How can we effectively leverage state-of-the-art AI technologies to build an accurate, interpretable, and non-invasive diagnostic model for early pancreatic cancer detection – in order to support doctors and save as many lives as possible?”

By addressing (and solving) this question, our study contributes to a growing body of research focused on combining cutting-edge AI tools with medical biomarker data to offer practical, real-world impact in the field of healthcare diagnostics.

**1.9 Objective**

The main objectives of the presented research can be summarized as follows:

* **Develop** an AI-based solution for early detection of pancreatic cancer using urinary biomarker data.
* **Utilize** state-of-the-art technologies in AI including 🡪 Machine Learning (ML), Deep Learning (DL), Foundation Models (TabPFN), and Explainable AI (XAI).
* **Deliver** a diagnostic model that combines high predictive accuracy with interpretability, allowing (and enabling) doctors to understand and trust the predictions.
* **Contribute** to the medical AI literature by benchmarking multiple imputation and classification techniques, with a particular focus on evaluating the potential of TabPFN in tabular healthcare data.

**1.10 Research Questions**

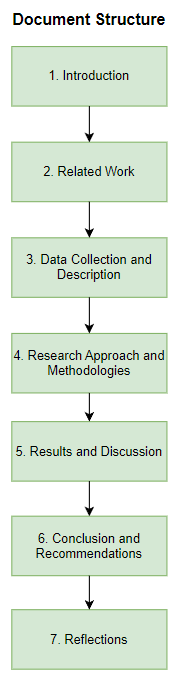
In this research, we aim to address the following key questions:

1. What is the most effective approach for handling missing values in urinary biomarker data for pancreatic cancer?
2. Which type of model achieves the best performance for pancreatic cancer detection: (a) traditional machine learning, (b) deep learning, or (c) foundation models like TabPFN?
3. Can TabPFN outperform conventional models in terms of accuracy and reliability on small medical datasets?
4. What is the most suitable classification model for accurate and early detection of pancreatic cancer?
5. Which hyperparameter tuning method is more effective: (a) GridSearchCV or (b) RandomizedSearchCV?
6. How can the final model be made interpretable and understandable for doctors using Explainable AI (XAI) techniques?
7. Is it possible to build a model that is both highly accurate and easily interpretable?
8. Can modern AI techniques help bridge existing gaps in the medical literature on pancreatic cancer detection – by providing a non-invasive, accurate, and interpretable diagnostic tool?

**1.11 Structure of the Document**

This document is structured (and split) into seven main sections to provide a comprehensive overview of the research process, methodologies, and findings, as illustrated in **Figure 13.**

* **Section 1 – Introduction:** Presents the background and context of the research, introduces key *core* concepts such as: AI, ML, DL, XAI, and TabPFN, and defines the problem, objectives, and research questions.
* **Section 2 – Related Work:** Reviews related literature, summarizing key methodologies, findings, and limitations of existing studies, and identifies gaps that this research aims to address (and solve).
* **Section 3 – Data Collection and Description:** Describes the dataset used in this study, including data sources, features, missing values, and pre-processing steps.
* **Section 4 – Research Approach and Methodologies:** Outlines the methodology used, including imputation techniques, classification models, hyperparameter tuning, and evaluation metrics.
* **Section 5 – Results and Discussion:** Presents the experimental results, compares model performances across different imputation strategies, and discusses the findings in relation to the research questions.
* **Section 6 – Conclusion and Recommendations:** Concludes the study by summarizing key findings, offering recommendations, and proposing potential future work.
* **Section 7 – Reflections:** Provides a reflective analysis of the research journey, focusing on the selected and alternative methodologies, lessons learned, and recommendations for future considerations (and improvements).

**Figure 13 – Research Document Structure**

# **2. Related Work**

The research papers reviewed in this section focus on two main areas: (1) **pancreatic cancer detection** and (2) **cutting-edge Artificial Intelligence (AI) techniques**, including machine learning (ML), deep learning (DL), foundation models, and explainable AI (XAI). Our aim (and goal) is to examine what has already been done in the field of pancreatic cancer detection, identify the latest advancements in AI, and explore how these technologies can be applied to improve early diagnosis. For each research paper, we provide a critical analysis of its methodology, key contributions, relevance to our research, limitations, and how our work aims to build upon and advance its findings further more.

**2.1 Research Paper 1:** “Accurate predictions on small data with a tabular foundation model” (Hollmann et al., 2025)

In this paper, the researchers introduce **TabPFN**, a novel tabular foundation model that brings the power of deep learning (DL) into the domain of tabular data – a field that has long been dominated by traditional machine learning models such as: **CatBoost**, **XGBoost (XGB)**, and **Random Forest (RF)**. While deep learning has shown remarkable success in areas like **natural language processing** (e.g., ChatGPT, BERT) and **computer vision (CV)**, it has struggled to outperform classical models on structured tabular datasets. This paper presents a new approach that aims to bridge and fill that gap.

What makes TabPFN unique is its **training paradigm**. Unlike traditional models that are trained on a single dataset, **TabPFN is pre-trained on millions of synthetic tabular datasets**. This massive pre-training phase allows it to generalize extremely well to new, unseen real-world tabular datasets in a **zero-shot** setting – meaning it does not require any additional training or fine-tuning before making predictions.

**Figure 14** illustrates the overall architecture of TabPFN.

* In **Figure 14a**, the model is trained on synthetic datasets where it learns to predict entire target vectors in a single forward pass.
* In **Figure 14b**, we see the internal structure of a TabPFN layer, which includes **1D feature attention**, **1D sample attention**, and a **multi-layer perceptron (MLP)** that generates the final prediction. This architecture allows the model to attend across both features and samples simultaneously, capturing complex relationships (and patterns) in the data.

**A diagram of a diagram

AI-generated content may be incorrect.Figure 14 – The TabPFN architecture. Source: Adapted from (Hollmann *et al.*, 2025)**

To better understand the innovation of TabPFN, **Figure 15** presents a comparison between TabPFN and traditional models like CatBoost and XGBoost (XGB) across key aspects.

**Figure 15 – Key Differences Between TabPFN and Traditional ML Models (e.g., CatBoost, XGBoost)**

|  |  |  |
| --- | --- | --- |
| Aspect | CatBoost, XGBoost (XGB), etc. | TabPFN |
| 1. Training Method | Trains on a single dataset. | Trained on millions of synthetic datasets (foundation model). |
| 2. Prediction Time | Fast. | Relatively slower (due to transformer-based architecture). |
| 3. Performance | Strong but may struggle with small data. | High accuracy on small data without further fine-tuning. |
| 4. Innovation | Uses boosting-based learning. | Uses transformer-based 2D attention over features and over samples. |

In our research, we propose to apply the **TabPFN model** for the task of **pancreatic cancer detection** using urinary biomarker data. To the best and latest of our knowledge, **no *any* prior study has evaluated TabPFN in the context of pancreatic cancer**. Thus, *therefore*, a key contribution of our study is to investigate how TabPFN compares to established machine learning (ML) models such as: **CatBoost**, **XGBoost (XGB)**, and **deep learning (DL) models**, in terms of predictive performance and interpretability.

By exploring this direction, we aim to determine whether TabPFN can offer a more accurate and robust diagnostic tool for early pancreatic cancer detection – especially given the small and structured nature of our dataset.

**2.2 Research Paper 2:** “A combination of urinary biomarker panel and PancRISK score for earlier detection of pancreatic cancer: A case–control study” (Debernardi et al., 2020)

This paper highlights and shows the urgent need for early detection of pancreatic cancer, which remains one of the most lethal cancers worldwide. The authors explain that the disease is often diagnosed at an advanced stage, contributing to a five-year survival rate of only around 9%. However, they emphasize that if pancreatic cancer is detected at an early stage—when tumors are still small and surgically resectable—the survival rate can increase dramatically to 🡪 60–70%. These findings underscore the life-saving potential of early diagnosis and reinforce the **importance of ongoing research in this area**.

In addition to its clinical contributions, this paper provides a high-quality, real-world dataset collected from multiple hospitals and clinical institutions. The dataset includes urinary biomarker data and other critical variables relevant to early pancreatic cancer detection. **In our research, we plan to use this dataset** as the foundation for developing and evaluating advanced AI-based diagnostic models. Its credibility, richness, and real-world origin make it especially valuable (and worthwhile) for building reliable and impactful machine learning (ML) solutions in the healthcare – medical domain.

**2.3 Research Paper 3:** “Automated classification of urine biomarkers to diagnose pancreatic cancer using 1-D convolutional neural networks” (Karar, El-Fishawy and Radad, 2023)

In this research paper, the authors developed and compared four different models for pancreatic cancer detection using biomarker data: **MLP (Multi-Layer Perceptron)**, **Random Forest (RF)**, **1D CNN**, and a custom-designed **1D CNN + LSTM** model. Their main contribution was the introduction of the 1D CNN + LSTM hybrid model, which achieved the highest performance among all the tested approaches. As shown in **Figure 16**, this model was able to classify most cases correctly, with only a few errors across all classes (1. Healthy, 2. Benign, 3. PDAC).

This study demonstrates two important lessons relevant to our research. First, it confirms that **AI models can accurately detect pancreatic cancer** using non-invasive urinary biomarkers. Second, it shows that **advanced architectures like 1D CNN + LSTM outperform both traditional machine learning (ML) models and simpler neural networks**. This insight supports the idea that more powerful or creative models – such as TabPFN in our case – can bring new potential to solving the early detection problem in pancreatic cancer.

**A screenshot of a chart

AI-generated content may be incorrect.Figure 16 – Confusion Matrices of the Four Models Evaluated in the Study. Source: Adapted from (Karar, El-Fishawy and Radad, 2023)**

**2.4 Research Paper 4:** “Can Everyday AI be Ethical? Machine Learning Algorithm Fairness” (Besse et al., 2018)

In this research paper, the authors emphasize the importance of Explainable AI (XAI) to avoid serious mistakes caused by black-box AI models. They give a well-known example, as shown in **Figure 17**, where a machine learning (ML) model misclassified a husky as a wolf – not because of the animal itself, but because it was trained on images where wolves were always in snowy backgrounds. The model learned to focus on snow instead of the animal itself, which led to completely wrong and invalid decisions.

**Figure 17 – Misclassification of a Husky as a Wolf Due to Snow Background. Source: Adapted from (Besse *et al.*, 2018)**

From this paper, we learn that explainability is not just helpful – it’s necessary, especially in high-stakes fields like 🡪 healthcare. In our research, we aim to apply XAI techniques to make sure our models for pancreatic cancer detection are learning for the right reasons, and not making *any* biased decisions based on irrelevant patterns like: (a) gender or (b) age. This helps us ensure (and make sure) that the model is fair, trustworthy, and medically reliable and robust.

**2.5 Research Paper 5:** “Early diagnosis of pancreatic cancer by machine learning methods using urine biomarker combinations” (ACER et al., 2023)

In the research paper by (ACER *et al.*, 2023), the authors used the same urinary biomarker dataset originally introduced by (Debernardi *et al.*, 2020), which we also use in our research. The authors applied seven different machine learning (ML) models to classify pancreatic cancer cases: Gradient Boosting Classifier (GBC), LightGBM, AdaBoost, Random Forest (RF), K-Nearest Neighbors (KNN), Naïve Bayes (NB), and Support Vector Machine (SVM). Among these, they found that **Gradient Boosting Classifier (GBC)** achieved the highest accuracy and performance, as shown in **Figure 18** below.

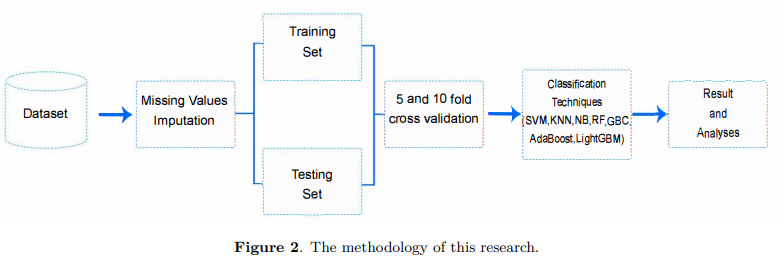
**Figure 18 – Performance Comparison of 7 ML Models Using 5-Fold and 10-Fold Cross-Validation. Source: Adapted from (ACER *et al.*, 2023)**

A table with numbers and a number of objects

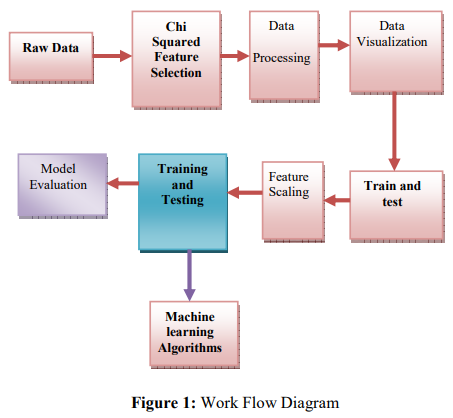
AI-generated content may be incorrect.

The methodology followed in their study is summarized in **Figure 19**, which outlines their process: starting from missing value imputation, to training/testing split, cross-validation (CV), and finally classification.

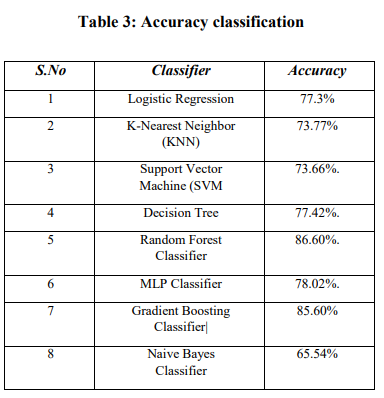
**Figure 19 – Methodology Pipeline Used in the Study by (ACER *et al.*, 2023)**



However, a major limitation and drawback of this paper is that the authors did **not apply any hyperparameter tuning techniques** such as: (1) **GridSearchCV** or (2) **RandomizedSearchCV**, which could significantly affect model performance and ranking. For example, with proper tuning, a model like LightGBM may outperform GBC. In our research, we aim to fill (and address) this gap by applying extensive hyperparameter tuning to ensure a more comprehensive and fair evaluation of all models, leading to more accurate and optimized results.

**2.6 Research Paper 6:** “A Comprehensive exploration of machine learning in early detection with a focus on lung and pancreatic cancer for revolutionizing cancer diagnostics” (Ravikumar et al., 2024)

In this research paper, the authors used the same urinary biomarker dataset originally introduced by (Debernardi *et al.*, 2020), which we also use in our research. They implemented eight different machine learning (ML) classification algorithms, including Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Decision Tree (DT), Random Forest (RF), MLP, Gradient Boosting, and Naïve Bayes (NB). Their findings showed that the **Random Forest (RF) Classifier** achieved the highest accuracy (86.6%) for the task of pancreatic cancer detection as illustrated in **Figure 20** below.

**Figure 20 – Accuracy Comparison of 8 ML Models for Pancreatic Cancer Detection. Source: Adapted from (Ravikumar *et al.*, 2024)**

The overall methodology followed in their study is summarized in **Figure 21** given below, showing steps such as feature selection using the Chi-squared test, data preprocessing, feature scaling, model training, and model evaluation.

**Figure 21 – Workflow of the Study Conducted by (Ravikumar *et al.*, 2024)**

However, we identify two major limitations (and drawbacks) in their study.

(1) First, the authors did **not apply any hyperparameter tuning** techniques like: (a) GridSearchCV or (b) RandomizedSearchCV, which could have significantly boosted the performance and fairness of their model comparisons.

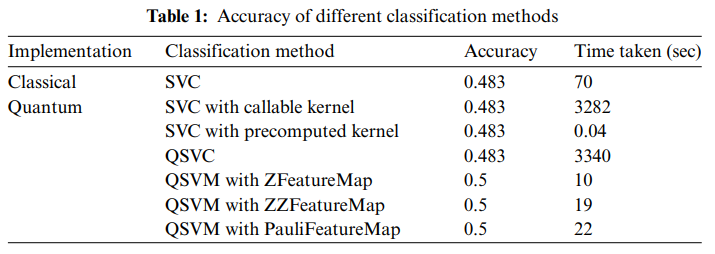
(2) Second, the **feature selection method they used introduced irrelevant features** into the training process – such as: `”Patient Cohort”` and `Sample Origin`. These features are not biologically meaningful (or relevant). `Patient Cohort` reflects whether the sample is from an old or new batch of data, and `Sample Origin` may reflect the geographical location where the sample was collected. Including these could introduce unwanted **bias** and may lead to **overfitting issues**, as the model could unintentionally learn from dataset artifacts rather than true disease-related patterns.

In our research, we aim to fill (and address) these gaps by applying **robust hyperparameter tuning techniques** to ensure a more comprehensive and accurate evaluation of model performance. Additionally, we use **carefully validated feature selection methods** that exclude misleading or irrelevant features, ensuring that our models learn from only the most reliable and meaningful biological indicators.

Lastly, while the authors recommend exploring more advanced machine learning (ML) approaches in future work, we go one step further by incorporating **deep learning (DL) architectures** and testing the **cutting-edge TabPFN foundation model**, which has never been explored before in this context. This positions our work as a forward-thinking contribution to the ongoing advancement of pancreatic cancer diagnostics.

**2.7 Research Paper 7:** “Pancreatic Cancer Data Classification with Quantum Machine Learning” (Saxena and Saxena, 2023)

Even going further, in the research paper by (Saxena and Saxena, 2023), the authors used the same urinary biomarker dataset introduced by (Debernardi *et al.*, 2020) to investigate whether quantum computing technology can bring better results and stronger predictive power than traditional machine learning (ML) models. Specifically, they compared the classical Support Vector Classifier (SVC) with six different quantum-based implementation approaches, as illustrated (and shown) in **Figure 22**.

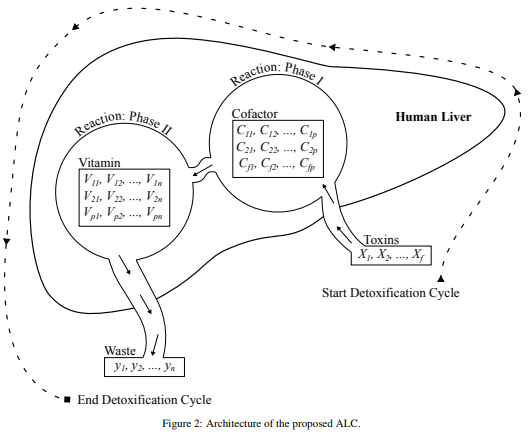
**Figure 22 – Accuracy and Time Comparison of Classical vs Quantum Models for Pancreatic Cancer Classification. Source: Adapted from (Saxena and Saxena, 2023)**

Their goal was to test whether quantum models like QSVC (Quantum Support Vector Classifier) could outperform classical *conventional* models in terms of classification accuracy and computational efficiency. From the results they obtained, we can conclude that while quantum computing shows **promising results in terms of time complexity and training speed**, it still struggles when it comes to predictive accuracy. The quantum implementations provided **significant speedups** – especially when using feature maps like ZFeatureMap and ZZFeatureMap – but the overall accuracy remained low, around **48.3%** to **50%**, which is still extremely weak compared to state-of-the-art machine learning models.

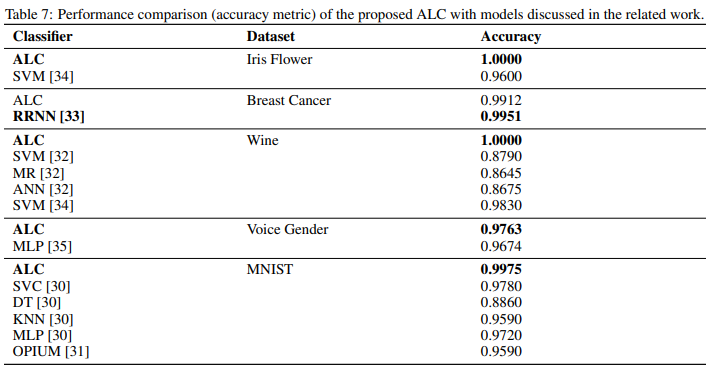
Thus, *therefore*, this study highlights a key trade-off: **quantum computing currently offers excellent computational efficiency**, but its **predictive performance is still limited** for small and structured medical datasets like ours.

**2.8 Research Paper 8:** “ARTIFICIAL LIVER CLASSIFIER: A NEW ALTERNATIVE TO CONVENTIONAL MACHINE LEARNING MODELS” (Jumaah, Ali and Rashid, 2025)

In this brand-new research paper, the authors introduce a completely novel machine learning model (ML) known as the **Artificial Liver Classifier (ALC)**. The motivation behind this model is inspired by the human liver – particularly its **detoxification process**. As shown in **Figure 23**, the architecture of ALC mimics how the liver processes and neutralizes harmful substances (toxins), using artificial equivalents called: (a) **Cofactors** and (b) **Vitamins**, which play a similar role in transforming inputs (toxins) into harmless outputs (waste). The entire classification process is modeled in two main phases – just like the biological liver’s detox cycle.

**Figure 23 – Architecture of the proposed Artificial Liver Classifier (ALC). Source: Adapted from (Jumaah, Ali and Rashid, 2025)**

The researchers conducted extensive experiments comparing ALC to traditional machine learning models (such as: SVM, KNN, MLP, and Decision Trees) across a variety of datasets including **Iris**, **Breast Cancer**, **Wine**, **Voice Gender**, and **MNIST**. As shown in **Figure 24**, ALC often outperformed or matched the best models in terms of classification accuracy – achieving **100% accuracy on Iris and Wine datasets**, and **99.75% accuracy on the MNIST dataset**. These results suggest that ALC has the potential to become a strong and competitive alternative to conventional models, especially in small and structured datasets.

**Figure 24 – Performance comparison of the proposed ALC model with other models. Source: Adapted from (Jumaah, Ali and Rashid, 2025)**

To our knowledge, no previous research has explored or evaluated the **ALC model in the context of pancreatic cancer disease detection**. Thus, *therefore*, a key contribution of our study is to investigate and compare the ALC model alongside other state-of-the-art machine learning, deep learning (DL), and foundation models. By doing so, we aim to assess its performance on a real-world medical dataset and determine whether this bio-inspired model can contribute effectively to early cancer diagnosis.

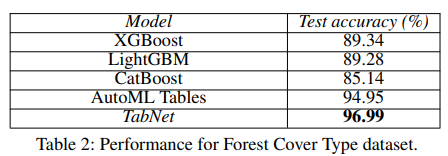
**2.9 Research Paper 9:** “TabNet: Attentive Interpretable Tabular Learning” (Arik and Pfister, 2021)

In this influential paper, the authors from Google AI introduce **TabNet**, a novel deep learning architecture designed specifically for **structured/tabular datasets**. Unlike traditional neural networks that treat all features equally, TabNet uses a **sparse attention mechanism** to selectively focus and concentrate on the most relevant features for each individual prediction. This attention-based feature selection is done in a sequential and interpretable manner, allowing TabNet to learn which features matter most at each given decision step.

This combination of **performance and interpretability** sets TabNet apart from other models. It not only delivers high predictive accuracy but also provides insights into which features were important for each prediction – making it a strong candidate for **high-stakes (and sensitive) fields like healthcare**.

The authors compared TabNet to popular machine learning (ML) models such as: **XGBoost (XGB)**, **LightGBM**, **CatBoost**, and **AutoML Tables** using the **Forest Cover Type** dataset. As shown in **Figure 25**, TabNet outperformed all other models, achieving a test accuracy of 96.99%, outperforming even Google’s AutoML system.

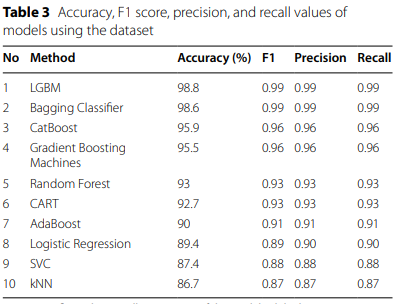
**Figure 25 – Test Accuracy comparison of TabNet and other models on the Forest Cover Type dataset. Source: Adapted from (Arik and Pfister, 2021)**



These results suggest that TabNet is not only accurate; but also generalizable across different tabular tasks. Therefore, **we believe TabNet is worth exploring in our pancreatic cancer detection pipeline**, especially given the tabular nature of our biomarker dataset and our strong emphasis on interpretability.

**2.10 Research Paper 10:** “Data privacy-aware machine learning approach in pancreatic cancer diagnosis” (AKMEŞE, 2024)

In the paper by (AKMEŞE, 2024), the authors applied **10 different machine learning (ML) algorithms** on the urinary biomarker dataset introduced by (Debernardi *et al.*, 2020). As shown in **Figure 26**, they found that the top-performing model was **LGBM**, achieving an impressive **accuracy of 98.8%** – the highest among all the models they tested.

**Figure 26 – Accuracy, F1 score, precision, and recall of 10 machine learning models. Source: Adapted from (AKMEŞE, 2024)**

To evaluate their results in a broader context, (AKMEŞE, 2024) compared the performance of their models with the results from other research papers in the field. As shown in **Figure 27**, their LGBM model was found to be the **best-performing model in the literature**, outperforming other studies that used urine biomarker datasets and models such as: GBC, 1D CNN-LSTM, and Random Forest (RF).

**Figure 27 – Comparison of classification accuracies from different research papers. Source: Adapted from (AKMEŞE, 2024)**

A screenshot of a data report

AI-generated content may be incorrect.

After carefully reviewing (and examining) their study, we found (and discovered) that **one of the main reasons they were able to achieve such high results** was due to their unique pipeline, which is clearly illustrated in **Figure 28**. In particular, they applied **outlier detection using the Interquartile Range (IQR)** method, specifically to the **plasma CA19\_9 biomarker**. Any outliers detected were **replaced with the mean value**, which likely helped stabilize the model training process and improve accuracy.

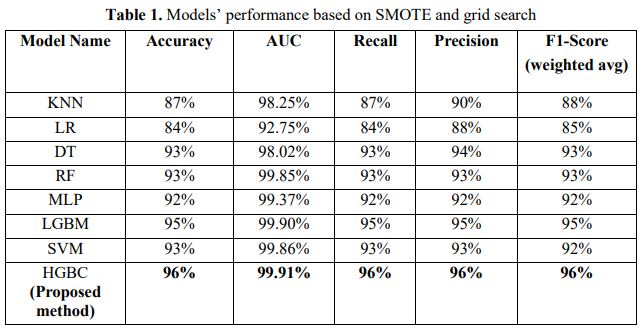
A diagram of a software system

AI-generated content may be incorrect.**Figure 28 – The proposed machine pipeline. Source: Adapted from (AKMEŞE, 2024)**

This approach highlights (and shows) the importance of **handling outliers properly** and following a well-structured pipeline. From this, we learn that applying **IQR-based outlier detection** and **careful processing steps** could be essential for improving performance in our own pancreatic cancer disease detection pipeline.

**2.11 Research Paper 11:** “Malicious URL Detection using optimized Hist Gradient Boosting Classifier based on grid search method” (Maftoun et al., 2024)

In the paper by (Maftoun *et al.*, 2024), the authors investigated the effectiveness of eight different machine learning (ML) algorithms for the task of malicious URL (Uniform Resource Locator) detection, as illustrated in **Figure 29**. A key contribution of their research was the introduction and evaluation of the **HistGradientBoostingClassifier (HGBC)** – a model that, to the best of their knowledge, had not been previously applied to this task. Unlike traditional models, HGBC showed superior performance across all evaluation metrics, achieving 96% accuracy, 99.91% AUC, and a 96% F1-score. This success was attributed to the model’s high efficiency, scalability, and the use of GridSearchCV for hyperparameter tuning. By demonstrating how HGBC can outperform other classical models when optimized correctly, the study highlighted (and showed) its strong potential in structured classification problems.

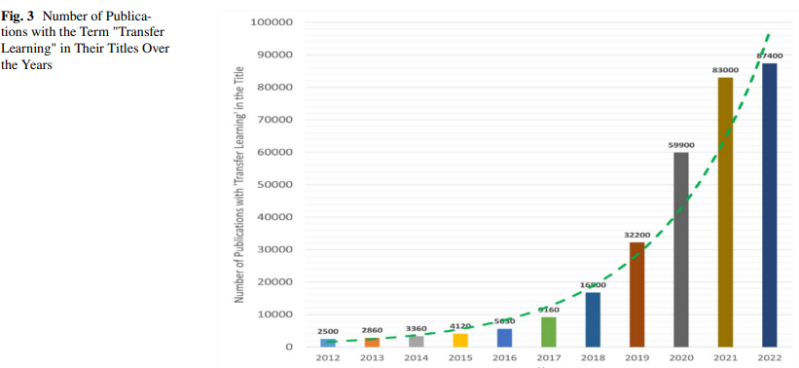
**Figure 29 – Performance Comparison of Eight ML Models for Malicious URL Detection. Source: Adapted from (Maftoun *et al.*, 2024)**

Learning from this paper, we believe that HGBC is a promising model worth exploring in our research on pancreatic cancer detection, especially since no prior work has evaluated it in this specific medical context. Given its performance in another structured dataset and its boosting-based design, it has the potential to serve as a powerful alternative to traditional – *conventional* ensemble models. We therefore include HGBC as one of the baseline models to be evaluated and compared within our experimental framework.

**2.12 Research Paper 12:** “A review of recent advances and strategies in transfer learning” (Gholizade et al., 2025)

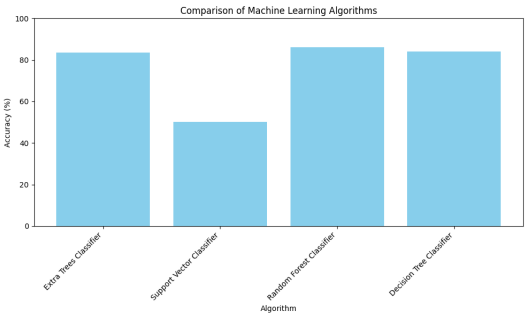
In their comprehensive review paper, (Gholizade *et al.*, 2025) explore the evolving landscape of **Transfer Learning (TL)** and present a detailed overview of its paradigms, methodologies, and real-world applications. The authors define transfer learning as a **technique that focuses on transferring knowledge from a source domain or task to a target domain or task**, especially useful (and helpful) when labeled data in the target domain is limited (and scarce).

The paper emphasizes the growing importance of transfer learning in the AI community, as shown in **Figure 30**, which illustrates a sharp increase in the number of academic publications that include the term “””transfer learning”””. This trend highlights the rising recognition of TL’s value in diverse fields such as: computer vision (CV), Natural Language Processing (NLP), and healthcare.

**Figure 30 – Number of Publications Containing the Term “”Transfer Learning””. Source: Adapted from (Gholizade *et al.*, 2025)**

Furthermore, **Figure 31** in the paper offers a clear visual explanation of the transfer learning process. The upper part of the figure represents the **source domain**, where a model is trained on a large dataset. This results in a **source model** with learned weights. In the lower part, the same model is then **fine-tuned on a smaller dataset** from the **target domain**, allowing it to adapt to the new context with less training data. This method significantly enhances **model generalization and efficiency**, particularly in data-scarce environments.

**Figure 31 – Schematic Diagram of the Transfer Learning Process. Source: Adapted from (Gholizade *et al.*, 2025)**

A diagram of a diagram of a computer

AI-generated content may be incorrect.

Learning from this work, we strongly believe (and think) that **transfer learning holds great promise for transforming cancer detection**, including the early diagnosis of pancreatic cancer. Since pancreatic cancer datasets are typically small, pretraining models on large-scale general medical datasets and subsequently fine-tuning them on domain-specific biomarker data could significantly enhance and improve predictive performance. This approach not only addresses data scarcity; but also enables better generalization to unseen patterns.

Although our current research does not directly implement transfer learning due to project scope and resource constraints, we consider it a compelling direction for **future work**. In particular, applying transfer learning to boost the accuracy and adaptability of AI-based pancreatic cancer diagnostics represents a promising research avenue that merits deeper investigation in upcoming *future* studies.

**2.13 Research Paper 13:** “Early Detection: Machine Learning Techniques in Pancreatic Cancer Diagnosis” (Sai et al., 2024)

In the study conducted and made by (Sai *et al.*, 2024), the authors used the urinary biomarker dataset originally introduced by (Debernardi *et al.*, 2020) and applied four traditional machine learning (ML) algorithms: Random Forest (RF), Extra Trees Classifier, Decision Tree (DT), and Support Vector Machines (SVM). As illustrated in **Figure 32 given below**, they found that the Random Forest (RF) Classifier achieved the highest accuracy, reaching 86.34%, followed closely by the Extra Trees and Decision Tree (DT) models.

**Figure 32 – Accuracy Comparison of Four Models for Pancreatic Cancer Detection. Source: Adapted from (Sai *et al.*, 2024)**

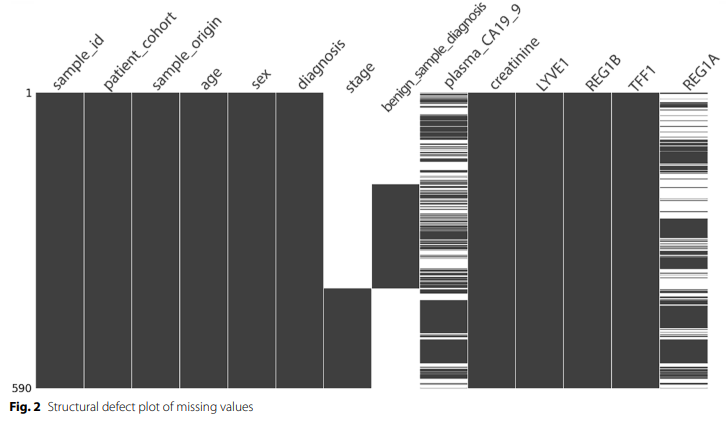
However, after carefully reviewing their methodology, we identified and found several major limitations in their study:

**(1) Lack of Hyperparameter Tuning:** The authors did not apply any optimization techniques such as: (a) GridSearchCV or (b) RandomizedSearchCV. Without tuning, models may fail to reach optimal performance – especially on complex biological datasets like this one.

**(2) Suboptimal Train-Test Split:** A 70-30 train-test split was used, leaving only 70% of the data for training. Given the small dataset size of 590 instances, this split may not allow models to fully learn the underlying patterns. A more effective split, such as 80-20, would provide more data for training while still reserving enough for evaluation.

**(3) Misleading Data Handling Decisions:** The authors claimed that columns such as `sample\_origin` and `patient\_cohort` had “excessive missing values” and removed them from the analysis. However, this claim is factually incorrect (and wrong). As shown in **Figure 33**, both features are fully complete with no missing values. Furthermore, these variables may contain predictive value (e.g., cohort differences, regional effects), and excluding them without proper investigation may negatively impact (and negatively affect) model performance.

**Figure 33 – Structural Defect Plot Showing that ‘sample\_origin’ and ‘patient\_cohort’ Contain No Missing Values. Source: Adapted from (AKMEŞE, 2024)**



By identifying and addressing such issues, our research aims to build a more accurate and trustworthy machine learning (ML) pipeline for early pancreatic cancer detection using urinary biomarkers.

**2.14 Research Paper 14:** “Enhanced CNN Model for Pancreatic Ductal Adenocarcinoma Classification Based on Proteomic Data” (Laxminarayanamma, Krishnaiah and Sammulal, 2022)

Also, similar to the paper by (Karar, El-Fishawy and Radad, 2023), in this research paper by (Laxminarayanamma, Krishnaiah and Sammulal, 2022), the authors developed (and created) a novel approach for the problem of pancreatic cancer detection using a 🡪 **Convolutional Neural Network (CNN)** model, as shown in **Figure 34**. The authors highlight their creative and innovative approach of applying CNN models – traditionally used for image data – to structured tabular data like biomarkers. They explain how the data was preprocessed and reshaped into a format that CNNs can ingest and learn from effectively (and efficiently).

**Figure 34 – CNN Architecture Proposed for PDAC Classification Based on Tabular Biomarker Data. Source: Adapted from (Laxminarayanamma, Krishnaiah and Sammulal, 2022)**

A diagram of a diagram

AI-generated content may be incorrect.

When they evaluated their CNN model and compared it with a traditional Artificial Neural Network (ANN) model, the results showed a clear performance advantage. Their CNN model achieved an accuracy of **95%**, while the ANN model only reached **70%**, as illustrated in **Figure 35**.

**Figure 35 – Accuracy Comparison Between ANN and CNN Models Over Training Epochs. Source: Adapted from (Laxminarayanamma, Krishnaiah and Sammulal, 2022)**

A graph showing the results of a graph

AI-generated content may be incorrect.

However, while these CNN models deliver high predictive power and impressive accuracy, a key limitation of this study is the lack of **explainability**. Despite the excellent performance of the CNN model, the authors did not explore or apply any **Explainable AI (XAI)** techniques to interpret their model’s decisions. This makes the model difficult to trust in high-stakes applications such as: cancer diagnosis. In contrast, our research aims to address (and solve) this gap by not only building high-performing models; but also ensuring interpretability and transparency using modern XAI (Explainable AI) techniques and technologies.

# **3. Data Collection and Description**

**3.1 Primary Data**

When it comes to primary data in my research, I designed and conducted a creative and insightful survey for the CRP (Computing A Research Project) course as a main source of primary data. The purpose (and aim) of this survey was to understand what university students at my university (HTU – Al Hussein Technical University) think and feel about Artificial Intelligence (AI), especially in the context of healthcare and critical decision-making. I successfully collected responses from 20 different students, carefully selecting five key questions that would capture their awareness, opinions, trust, and attitudes toward AI. I intentionally limited the survey to five questions because university students are often busy, and I wanted to make sure the survey was quick and convenient for them to complete (and finish).

The five questions in my survey were:

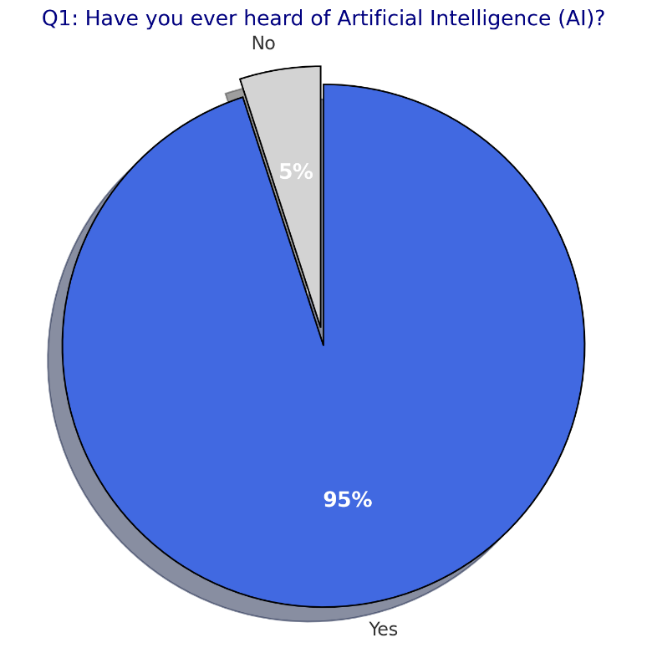
1. **Have you ever heard of Artificial Intelligence (AI)?**
2. **How would you describe your overall opinion of AI?**
3. **Which statement best matches your opinion about AI and jobs?**
4. **Do you think using AI in the healthcare field is beneficial?**
5. **If an AI-powered robot could perform surgery with 100% accuracy, how confident would you feel about undergoing such a surgery?**

Below, you can see screenshots of the actual survey questions and their summarized results as shown in Google Forms platform:



A screenshot of a survey

AI-generated content may be incorrect.



A screenshot of a survey

AI-generated content may be incorrect.

A pie chart with different colored circles

AI-generated content may be incorrect.

A blue rectangle with white stripes

AI-generated content may be incorrect.

**Analysis of Survey Results:-**

When analyzing (and studying) the survey, the results I got were really interesting and gave me valuable (and worthwhile) insights. Firstly, for the **first question**, an impressive **95% of the students** who filled out the survey said that they have heard about AI. This is amazing because it shows that almost all students are aware of this trending and modern technology, and only 5% haven’t heard about it. This suggests that AI has become a buzzword in society, especially due to its frequent appearance in 🡪 (a) television, (b) news, and (c) social media applications.

For the **second question**, I found that **80% of the students** think that AI is a good technology and a positive thing, while only 10% think that it is bad, and another 10% are neutral or unsure. This tells us that most students generally believe that AI (Artificial Intelligence) is a beneficial technology for humans, and they see its pros (advantages) as *far* outweighing and surpassing its cons (disadvantages).

A pie chart with a green circle and red and blue text

AI-generated content may be incorrect.Moving to the **third question**, the majority—**80% of students**—believe that AI will work (side-by-side) with humans and will not completely replace them. Only 10% think that AI will fully replace humans, while another 10% think that AI will never replace humans at all. This indicates and shows that most students have a balanced and realistic view of AI, expecting collaboration between both: (1) humans and (2) AI – *Artificial Intelligence* – in the future. It also shows that they are generally optimistic and excited about working with AI rather than being afraid (and anxious) of it.

A graph of different colored squares

AI-generated content may be incorrect.

A graph with green lines

AI-generated content may be incorrect.For the **fourth question**, a huge majority—**90% of students**—believe that AI is beneficial (and advantageous) for the healthcare domain and see positive applications for AI in medicine. Only 5% disagreed, and another 5% were unsure. This result demonstrates (and shows) strong trust in AI’s potential to help in sensitive and important fields like 🡪 healthcare domain.

However, the **last question** about confidence in an AI robot performing surgery with 100% accuracy had more mixed responses. The results were quite evenly distributed: some students were 100% confident, some were 80% confident but still had concerns, others were somewhat afraid and would prefer a human doctor, and the rest said they would never trust robots with such a critical (and sensitive) task. This shows that, while students generally see AI – *Artificial Intelligence* – as a good and useful technology (especially in healthcare), there is still some hesitation and lack of full trust in fully autonomous AI decisions—especially when it comes to life-and-death situations like 🡪 performing (and executing) a critical surgery.

To sum up, this survey was really beneficial and provided amazing insights into how university students (at HTU – AlHussein Technical University) think about AI. The results clearly show that most students are aware of AI, think it is a good thing, and see its benefits, especially in the healthcare domain. However, a large portion of students still do not fully trust AI to make critical decisions on its own. This highlights (and shows) the importance of our research focus on explainable AI (XAI), because making AI decisions transparent and understandable to humans can help build trust. It also supports the direction of our project—using AI in pancreatic cancer detection—because it shows that students are open to and supportive of AI (Artificial Intelligence) applications in healthcare, but also expect AI systems to be 🡪 (a) *fully* explainable, (b) reliable, and (c) *also* robust.

**Experience in Collecting Primary Data:-**

**My experience in collecting this primary data had both positives (pros) and challenges (cons).** One of the main difficulties I faced was finding enough students willing to fill out the survey—university students are often very busy, and some were reluctant to participate. It was also a challenge to design survey questions that were both meaningful and brief (*at the same time*). I had to carefully select just five questions because longer surveys would have discouraged participation. Ensuring that each question captured useful and unbiased information also took time and several *various* revisions.

**Merits and Limits of Primary Data:-**

**The main benefit of collecting primary data** is that I had full control over the kind of information I wanted to gather. I could design the survey to directly match the focus of my research and ask questions that addressed the exact gaps or topics I was interested in. This allowed (and enabled) me to get (a) fresh, (b) specific, and (c) directly relevant data tailored for my project.

**However, primary data collection also comes with certain limitations.** It takes much more effort and time to prepare the survey, find participants, and encourage them to respond—especially compared to using existing secondary data. The small sample size (20 students) also means the results may not be fully generalizable to all students or other populations. Finally, I was limited in how many questions I could reasonably ask, which may have restricted the depth of information I could gather (and accumulate).

**3.2 Secondary Data**

When it comes to secondary data, I used a real-world medical dataset that was originally published and made available by (Debernardi *et al.*, 2020). This dataset was chosen specifically because it comes from a credible and highly respected source in the medical research field. It provides real-world, high-quality urinary biomarker data for pancreatic cancer detection, which is both rare and valuable for research purposes. One of the main reasons for choosing this dataset is that it studies one of the most aggressive and lethal types of cancer – **pancreatic cancer** – making it directly relevant to the goals (and objectives) of our research.

As we can see in the image given below, the dataset contains **590 samples** (rows) and **14 features** (columns), which makes it a relatively small dataset. This is actually expected in the healthcare domain, since medical data is often hard to collect due to privacy concerns, strict regulations, and the need for careful patient consent.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.In the next visual, we can see a snapshot of the first two rows of the data. Each row corresponds to a unique patient sample, and each column captures different information about that sample, such as: age, sex, creatinine levels, REG1B levels, TFF1 levels, and so on. The key target variable is diagnosis, which represents whether the patient is (1) healthy, (2) has a benign condition, or (3) has pancreatic cancer.

To better understand the relationships between features, the correlation heatmap drawing below visualizes the correlation between the main biomarkers and the diagnosis column. We can observe that several variables, such as: LYVE1, REG1B, TFF1, and plasma\_CA19\_9, are quite strongly correlated with the diagnosis, highlighting their importance in detecting pancreatic cancer.

A graph with different colored squares

AI-generated content may be incorrect.

A graph of a patient's status

AI-generated content may be incorrect.Also, the bar chart below demonstrates the class balance in our data. The target diagnosis is evenly distributed between the three classes: Healthy (183 samples), Benign Condition (208 samples), and Pancreatic Cancer (199 samples). This is a positive aspect, as it means there is no major class imbalance that could bias the models, and we do not need to apply techniques like 🡪 SMOTE (Synthetic Minority Over-Sampling Technique) or *any* oversampling.

In-summary, our secondary dataset is from a credible medical research source, contains 590 samples and 14 features, and targets early pancreatic cancer disease detection. The dataset is balanced across all classes, with three categories: (1) healthy, (2) benign, and (3) pancreatic cancer.

**Merits and Limits of Secondary Data:-**

One of the biggest advantages (and benefits) of using secondary data in my research is the accessibility and convenience it provides, especially in the healthcare domain. Since medical data is notoriously difficult (and hard) to collect due to privacy laws, ethical restrictions, and the need for specialized equipment and consent, being able to use a well-established and openly available dataset from (Debernardi *et al.*, 2020) allowed me to focus directly on data analysis and model development rather than spending months (or even years) on data collection. This made my research much more feasible and efficient. Additionally, because the dataset was produced by a team of medical professionals and published in a reputable journal, I had a high level of confidence in its credibility and scientific accuracy—a crucial factor when dealing with something as sensitive as pancreatic cancer disease diagnosis.

Another important merit is cost-effectiveness. Instead of dedicating significant resources to creating a new dataset—which would be nearly impossible as a university student without access to clinical labs—I could work with high-quality real-world data at no extra cost. This also meant I had access to data that was broader and more diverse than anything I could have gathered myself. For example, the dataset includes samples from multiple patient cohorts and hospitals, which improves the representativeness and generalizability of my findings.

A diagram of a diagram

AI-generated content may be incorrect.But, *however*, using secondary data also comes with certain limitations (and constraints). Because I did not collect the data myself, I had no control over which features were included, how they were measured, or the protocols used for data collection. There are always concerns about possible biases in the way the original study was conducted—for instance, there may be demographic or geographical factors that are not obvious from the dataset alone. In some cases, I also found that certain clinical details I would have liked to analyze were simply not available, and there was no way to go back and ask for more information. Another limitation is that the data, while recent, may not include the latest biomarkers or reflect very recent advances in the field. As a result, there’s always a small risk that some findings could become outdated over time (*throughout time*).

Despite these challenges (and difficulties), the secondary data I used in this project gave me an invaluable foundation to explore advanced AI and machine learning (ML) models in the context of pancreatic cancer detection. The strengths—especially accessibility, credibility, and diversity—far outweighed the drawbacks for my research goals. In combination with my primary data (the survey I conducted), this allowed (and enabled) me to bring together both firsthand perceptions and real clinical evidence in one comprehensive project.

# **4. Onion Research Model**

When it comes to the **Onion research model**, it is simply – *in very simple words* – a structured research framework designed (and made) to help (and guide) researchers 🡪 (a) plan, (b) design, *and* (c) explain their research methodology in a step-by-step manner. It consists of six layered stages, starting from the outermost layer (**Philosophy**) and moving inward toward **Techniques and procedures**, much like peeling an onion vegetable layer-by-layer. Each layer represents a different aspect of research planning, guiding the researcher to make consistent, logical, and well-justified methodological choices.

Image Source 🡪 [Towards an Explicit Research Methodology: Adapting Research Onion Model for Futures Studies \* Journal of Futures Studies](https://jfsdigital.org/articles-and-essays/2018-2/towards-an-explicit-research-methodology-adapting-research-onion-model-for-futures-studies/)

In my research paper on “””Pancreatic Cancer Detection using Machine Learning, Deep Learning, TabPFN, and Explainable AI”””, I followed the Onion Research Model to structure my methodology clearly and professionally. Here’s how I applied (and utilized) each layer – specifically:

**1. Research Philosophy – Positivism**:

I adopted a **positivist philosophy** because my research relied on objective, measurable data from urinary biomarkers, and I focused on using statistical evaluations and numerical model results (such as 🡪 1. accuracy, 2. precision, 3. recall, and 4. F1-Score) to determine effectiveness. This reflects the belief that truth can be discovered (and found) through scientific, data-driven analysis.

**2. Approach to Theory Development – Deduction**:

My approach was **deductive**, meaning I started from established theories and models in the field of AI – *Artificial Intelligence* – and healthcare (such as: Random Forest, KNN, Logistic Regression, SVM, Decision Tree, TabPFN), and then tested them on real-world data to validate their performance for the task of pancreatic cancer disease detection.

**3. Methodological Choice – Mono Method Quantitative**:

I used a **mono method quantitative approach** because my research focused (and concentrated) entirely on numerical and statistical analysis. I did not collect interviews or *any* qualitative feedback. Instead, I trained and tested different AI models and measured their performance using quantifiable – measurable metrics.

**4. Strategy – Experimental and Data Analysis (Analytical)**:

I followed an experimental strategy, where I applied different data preprocessing, imputation, and modeling techniques on the dataset. This allowed (and enabled) me to run and conduct various multiple experiments and analyze which AI – *Artificial Intelligence* – approach (ML, DL, or TabPFN) gave the best (and top) results.

**5. Time Horizon – Cross-Sectional**:

I used a cross-sectional time horizon, meaning I analyzed the dataset as a snapshot at a single point in time. The data was already collected and did not span across multiple time periods or across multiple stages.

**6. Techniques and Procedures – Imputation, Modeling, Tuning, and XAI (Explainable AI):**

At the core of my research were the techniques and procedures used. These included applying 11 different imputation techniques, training and evaluating 26 classification models, performing hyperparameter tuning using: (a) GridSearchCV and (b) RandomizedSearchCV, and using Explainable AI (XAI) tools like SHAP (and LIME) to interpret the results and explain model decisions.

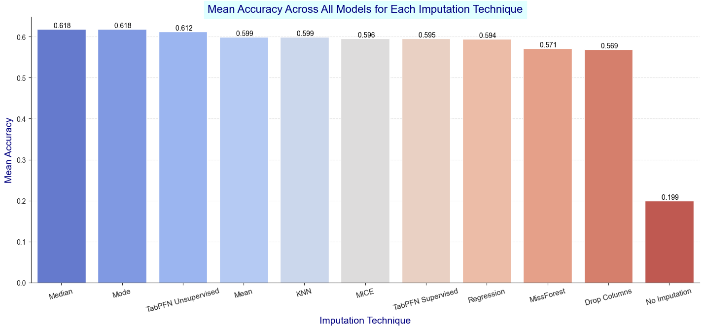
Overall, *generally*, applying the Onion Research Model was extremely beneficial (and really helpful) in guiding and organizing my research. It helped me define my objectives clearly, choose the right methods at every stage, and maintain a logical and structured flow from start to finish. Following this model ensured that my research was 🡪 (1) well-planned, (2) transparent, and (3) scientifically valid, which is especially important (and paramount) in a critical – and vital field like: pancreatic cancer disease detection.

# **5. Research Methodology**

For my research methodology, I followed a clear and systematic workflow consisting of a series of structured steps, as illustrated (and shown) in the flowchart given below. The process began with data collection, where I used both (1) **primary** and (2) **secondary** data sources. For primary data, I designed and conducted a survey among 20 university students at HTU (Al Hussein Technical University) [Home | HTU](https://www.htu.edu.jo/) to gather their insights and perceptions about Artificial Intelligence (AI), especially in healthcare. I analyzed these survey results to obtain valuable first-hand perspectives and highlight the importance of trust and explainability in AI.

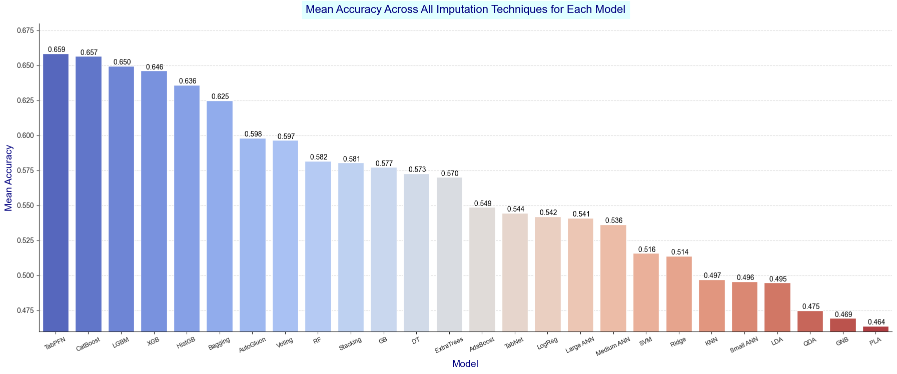
A diagram of a process

AI-generated content may be incorrect.



For secondary data, I adopted a more comprehensive and technical pipeline, focusing on the real-world urinary biomarker dataset for pancreatic cancer disease detection. The main (and core) steps in my methodology included:

**5.1 Data Understanding & Preparation**

A graph of a diagram

AI-generated content may be incorrect.I started with Exploratory Data Analysis (EDA), including visualizations and dimensionality reduction (such as: PCA, t-SNE, and UMAP) to understand the structure, separability, and characteristics of the data. As shown in the figure given below, these visualizations revealed how the different diagnostic groups (1. healthy, 2. benign, 3. pancreatic cancer) are distributed, with certain clusters being more distinct than other clusters.

**5.2 Missing Values Imputation**

Next, I tackled missing values by testing **11 different imputation techniques** (including Median, Mode, Mean, KNN, MissForest, MICE, Regression-based imputation, and TabPFN-based imputation). The performance of each technique was compared across multiple machine learning (ML) models. As shown in the accuracy comparison chart, Median and Mode imputation gave the best and most stable results, so I selected **Mode imputation** for subsequent *further* modeling steps.

**5.3 Model Development & Selection**

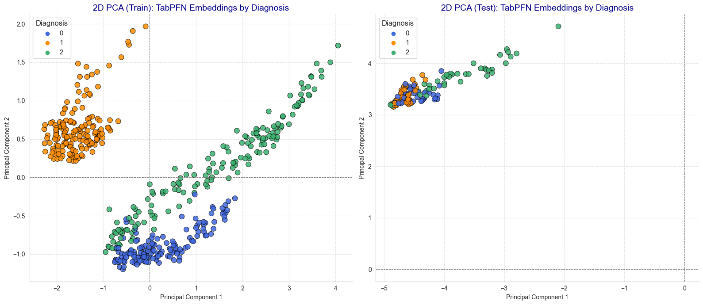
In the model selection phase, I evaluated over **26 different models** (traditional ML, deep learning (DL), and the new TabPFN model). The results indicated that TabPFN consistently achieved the highest performance—outperforming all other models tested. This finding demonstrates the strength of foundation models for small, structured healthcare datasets, making TabPFN my top choice for pancreatic cancer disease detection.

A screenshot of a computer generated image

AI-generated content may be incorrect.**5.4 Model Evaluation**

A chart with numbers and labels

AI-generated content may be incorrect.For evaluation, I used multiple metrics—**accuracy, precision, recall, and F1-score**—and analyzed confusion matrices to interpret the model’s predictive abilities. As seen in the confusion matrix given below, the TabPFN model performed well, accurately distinguishing between: (1) healthy, (2) benign, and (3) pancreatic cancer cases.

**5.5 Data Augmentation Using TabPFN**

To further enhance (and improve) model performance and address data scarcity, I explored **TabPFN-based data augmentation** techniques using the tabpfn-extensions package. The results were promising: all four TabPFN augmentation methods led to improved model accuracy. For example, the unsupervised + pseudo-labels method increased model accuracy by more than 2%. This novel step demonstrates (and shows) the practical value of data-centric AI strategies in healthcare research.

**5.6 TabPFN Embeddings**

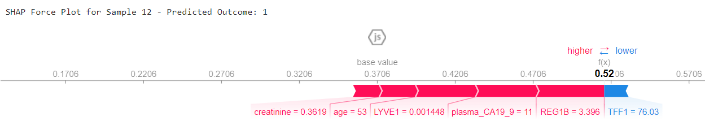
I also extracted **TabPFN embeddings** and visualized them using (1) PCA – *Principal Components Analysis*, (2) t-SNE (t-distributed Stochastic Neighbor Embedding), and (3) UMAP (*Uniform Manifold Approximation and Projection*) to better understand the feature representations learned by the model. These visualizations provided strong insights—for example, they revealed that the benign samples formed a separate, easily distinguishable cluster, while healthy and cancer groups were closer and harder to separate. This highlights the model’s interpretability and the clinical complexity of the data.

**5.7 Explainable AI (XAI)**

Recognizing the importance of interpretability, I applied XAI techniques—such as permutation importance and SHAP values—to explain the TabPFN model’s predictions. These tools revealed which biomarkers were most influential in the diagnostic process (e.g., plasma CA19-9, LYVE1, TFF1), and provided both global and individual-level interpretability. This supports the goal of making AI tools more transparent and trustworthy for clinical use.

A graph showing different colored bars

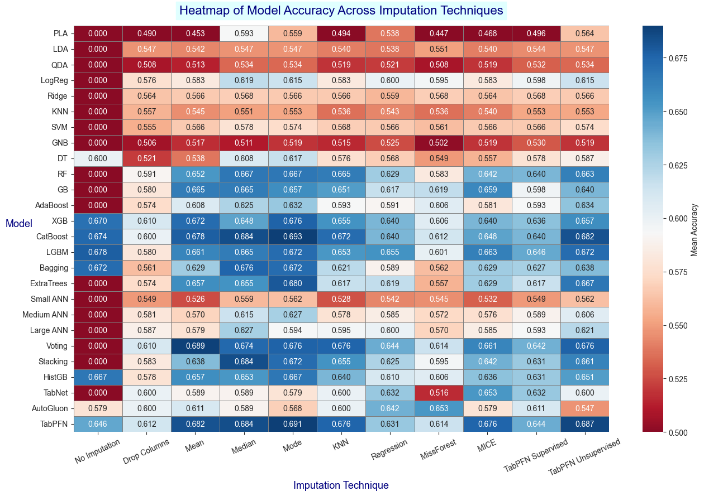
AI-generated content may be incorrect.



**5.8 Web Interface Deployment**

A screenshot of a medical test

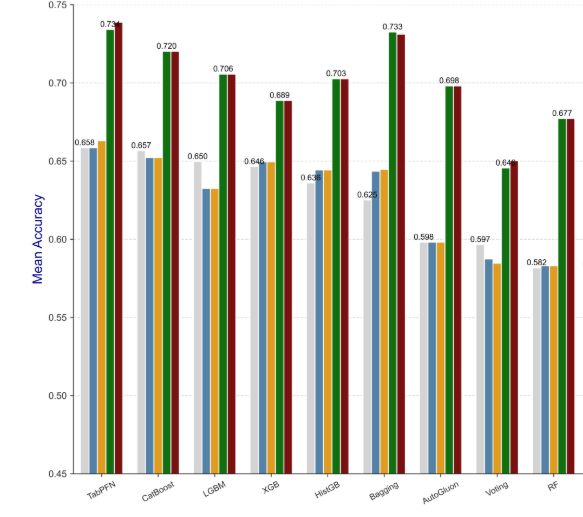
AI-generated content may be incorrect.Finally, at the end, to maximize the practical impact of my work, I deployed the best TabPFN model into a user-friendly **web interface application**. This web tool allows clinicians or researchers to input patient biomarker values and receive immediate, interpretable risk predictions for pancreatic cancer disease detection.

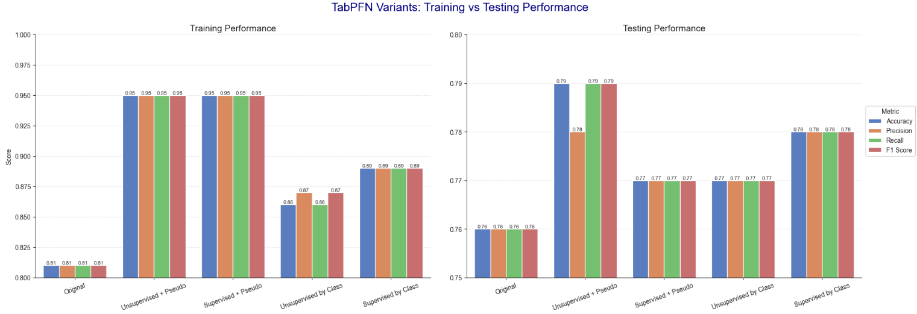


In-summary, my research methodology is comprehensive and innovative, covering every step from data collection (primary and secondary) through model development, model evaluation, augmentation, interpretability, and deployment. Key novel contributions include using TabPFN for pancreatic cancer disease detection, pioneering TabPFN-based data augmentation, extracting and visualizing TabPFN embeddings, and applying explainable AI (XAI) tools to improve transparency. This holistic approach demonstrates (and illustrates) how modern AI methods, combined with rigorous methodology, can advance early cancer detection and support real-world healthcare needs.

# **6. Results and Discussion**

For the results obtained in this research, the findings clearly demonstrate (and illustrate) that the **TabPFN model** was the best-performing model out of all 26 models tested, consistently achieving the highest and most stable results across a variety of imputation techniques. As shown in the correlation heatmap visualization given below, TabPFN outperformed both traditional machine learning (ML) and deep learning (DL) models, regardless of which imputation method was used. This highlights the remarkable robustness and adaptability of TabPFN, especially in small, structured medical datasets like ours.

In addition, we carefully measured and compared the **accuracy, precision, recall, and F1-score** for every model. TabPFN not only led in accuracy but also delivered top results in all other evaluation metrics, confirming its position as the most reliable and accurate model for early pancreatic cancer detection in our experiments. This directly addresses and meets one of our key research objectives: to benchmark the performance of TabPFN against a comprehensive set of alternative models, and to validate its suitability for medical diagnosis.

A chart with numbers and labels

AI-generated content may be incorrect.To further interpret and validate the model’s predictions, we analyzed **confusion matrices** and compared the performance on both (a) training data and (b) testing data. As seen in the confusion matrix and classification report below, TabPFN delivered balanced and consistent results, with high recall for pancreatic cancer cases and no signs of overfitting. The training and testing metrics—accuracy, precision, recall, and F1—were all around 0.76, reflecting strong generalization capability.

A number of numbers in a row

AI-generated content may be incorrect.A screenshot of a graph

AI-generated content may be incorrect.

Also, another important contribution of our research was the **testing and evaluation of TabPFN-based data augmentation**. By using advanced augmentation techniques available in the tabpfn-extensions package, we demonstrated that TabPFN not only performs well on original data but also benefits significantly from synthetic data generation. As shown in the performance charts and confusion matrices below, models trained with TabPFN augmentation consistently achieved **higher accuracy, precision, recall, and higher F1-scores** compared to models trained without augmentation. For example, the ““”unsupervised + pseudo labels”” method led to a noticeable improvement in classification metrics—raising the accuracy by more than 2% percent.

A screenshot of a computer generated image

AI-generated content may be incorrect.

# **7. Conclusion and Recommendations**

The aim (and core objective) of our study was to explore multiple goals, with the main focus being to evaluate the performance of the TabPFN model and see if it could outperform traditional machine learning (ML) and deep learning (DL) models in the task of pancreatic cancer disease detection. Our assumption was proven correct—**TabPFN**, with its groundbreaking transformer-based architecture, **delivered noticeably better results than any of the standard models we tested.** Because of this, we strongly recommend (and suggest) that researchers working with small or structured datasets, especially in healthcare, consider including TabPFN in their modeling toolkit, as it showed exceptional accuracy and robustness in our experiments.

Another important contribution of our paper was testing whether **TabPFN-based data augmentation** could further improve model performance. After running multiple *various* experiments with the data augmentation methods provided in the tabpfn-extensions package, we found clear evidence that these techniques boosted results—models trained with TabPFN augmentation achieved higher accuracy, precision, recall, and higher F1-scores. This finding is very promising, especially for researchers who face the common problem of limited data, as TabPFN data augmentation offers a practical and effective solution for enhancing model outcomes in small datasets.

We also introduced the **extraction and visualization of TabPFN embeddings** in our research. The results we obtained from visualizing these embeddings using (1) PCA – *Principal Components Analysis*, (2) t-SNE, and (3) UMAP (*Uniform Manifold Approximation and Projection*) were really impressive. These visuals provided valuable insights into the structure and separability of the data, making it easier for doctors and researchers to understand the relationships between different patient groups and biomarker features. We believe that exploring these embeddings can help clinicians gain deeper knowledge and even assist in real-world diagnostic decision-making.

Based on these findings, our key recommendations (and suggestions) are as follows. (1) First, **we recommend using the TabPFN model for small- to medium-sized tabular datasets**, especially in medical domains, because of its superior (and robust) performance. (2) Second, we encourage researchers to try TabPFN data augmentation if they are working with limited data and want to improve model accuracy and reliability. (3) Lastly, we suggest that others take advantage of TabPFN embeddings and visualization tools, as these can reveal important patterns and help interpret results in a more intuitive way.

For **future work**, there are several directions to consider. In our study, we evaluated TabPFN on only one dataset; it would be valuable to test the model on more datasets to further prove its generalizability and effectiveness. Due to limited computational resources, we could not perform exhaustive hyperparameter tuning, so using more powerful hardware or *even* cloud computing for deeper optimization could be a major improvement. Additionally, we recommend including more evaluation metrics—such as 🡪 ROC (*Receiver Operating Characteristic*) curve and AUC (*Area Under the Curve*) score—instead of focusing only on: (a) accuracy, (b) precision, (c) recall, and (d) F1-score. Future work could also experiment with automated feature selection techniques, which might further improve results. Finally, we only used a few explainable AI (XAI) tools in our project; expanding this by adding more visuals and using other libraries like LIME alongside SHAP would enhance the interpretability and trustworthiness of the model even more.

In-summary, our research highlights (and shows) the practical value of TabPFN for pancreatic cancer disease detection, demonstrates the benefits of novel augmentation and embedding techniques, and provides actionable insights and recommendations for future studies in the field.

# **8. Reflections**

After finishing (and completing) the research paper and looking back on my research journey, I can honestly say that it was an overall amazing and rewarding experience. I learned a lot of new things, both academically and personally. Throughout this project, I developed important skills—like how to search for and read research papers, summarize key findings, and write a proper – *appropriate* literature review. I also learned how to design a survey, collect responses, and analyze the results to extract meaningful insights. On top of that, I significantly improved my coding skills in Python3 programming language and learned how to work with new libraries, especially when it comes to implementing machine learning (ML) and deep learning (DL) models. Most importantly, I proved (and showed) to myself that I could complete a complex research project and show that TabPFN is a powerful and promising model for real-world problems.

However, this journey was not without its mistakes and pitfalls. One of the main issues I faced was the lack of access to advanced computational resources. All of my coding was done on my personal computer without any GPU (*Graphical Processing Unit*) acceleration or cloud computing. This made the training process for some models, especially TabPFN, extremely slow—sometimes taking more than five hours just to run a single experiment. In the future, I definitely plan to use cloud computing or more powerful computational hardware to speed up my work and allow for more ambitious (and aspiring) experiments.

Another limitation was with the primary data collection. I was only able to collect 20 survey responses because I had just two days to distribute the survey and get answers. If I had more time and could reach a larger group—like 50 or even 100 students— I think the results would have been much deeper and more representative. Also, I relied solely on a survey for primary data and did not conduct any interviews. Looking back, I realize that interviews with experts (such as: doctors or AI researchers or professionals) would have given me more qualitative data and could have led to even better and richer – *deeper* insights.

A further mistake was not spending enough (and sufficient) time on hyperparameter tuning for the AI – *Artificial Intelligence* – models. Because I was limited by time and computing power, I often chose quick results over more comprehensive (and thorough) optimization. If I had been more patient and allowed for longer tuning, I might have been able to get even better (and higher) results, especially with TabPFN and the deep learning (DL) models.

I also recognize that my literature review was not as broad (and wide) as it could have been. I read only 14 research papers—if I had read more, I could have developed a more innovative research idea and gained a wider – *deeper* understanding of artificial intelligence (AI) and its applications in the healthcare field – *domain*.

Despite these shortcomings (and weaknesses), the positive outcomes (pros) of the research journey far outweigh the negatives (cons). I learned how to write a professional research paper, cite sources correctly, and use tools like: Google Scholar ([Google Scholar](https://scholar.google.com/)) to find high-quality academic references. I also gained confidence in presenting my work, organizing my thoughts, and communicating complex ideas in a clear and structured way. Overall, I am grateful (and thankful) for the experience and believe that I have grown as a 🡪 (a) researcher, (b) student, and even as a (c) problem-solver.

# **References**

ACER, İ. *et al.* (2023) ‘Early diagnosis of pancreatic cancer by machine learning methods using urine biomarker combinations’, *Turkish Journal of Electrical Engineering and Computer Sciences*, 31(1), pp. 112–125. Available at: https://doi.org/10.55730/1300-0632.3974.

Adadi, A. and Berrada, M. (2018) ‘Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)’, *IEEE Access*, 6, pp. 52138–52160. Available at: https://doi.org/10.1109/ACCESS.2018.2870052.

*AI VS Traditional Programming - What’s the Difference?* (2024) *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/what-is-artificial-intelligence-ai-and-how-does-it-differ-from-traditional-programming/ (Accessed: 9 March 2025).

AKMEŞE, Ö.F. (2024) ‘Data privacy-aware machine learning approach in pancreatic cancer diagnosis’, *BMC Medical Informatics and Decision Making*, 24(1), p. 248. Available at: https://doi.org/10.1186/s12911-024-02657-2.

Arik, S.Ö. and Pfister, T. (2021) ‘TabNet: Attentive Interpretable Tabular Learning’, *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8), pp. 6679–6687. Available at: https://doi.org/10.1609/aaai.v35i8.16826.

Besse, P. *et al.* (2018) ‘Can Everyday AI be Ethical? Machine Learning Algorithm Fairness’. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=3391288 (Accessed: 4 April 2025).

Danielsson, J. and Uthemann, A. (2024) ‘On the use of artificial intelligence in financial regulations and the impact on financial stability’. arXiv. Available at: https://doi.org/10.48550/arXiv.2310.11293.

Debernardi, S. *et al.* (2020) ‘A combination of urinary biomarker panel and PancRISK score for earlier detection of pancreatic cancer: A case–control study’, *PLOS Medicine*, 17(12), p. e1003489. Available at: https://doi.org/10.1371/journal.pmed.1003489.

Esteva, A. *et al.* (2019) ‘A guide to deep learning in healthcare’, *Nature Medicine*, 25(1), pp. 24–29. Available at: https://doi.org/10.1038/s41591-018-0316-z.

Gholizade, M. *et al.* (2025) ‘A review of recent advances and strategies in transfer learning’, *International Journal of System Assurance Engineering and Management*, 16(3), pp. 1123–1162. Available at: https://doi.org/10.1007/s13198-024-02684-2.

Hollmann, N. *et al.* (2025) ‘Accurate predictions on small data with a tabular foundation model’, *Nature*, 637(8045), pp. 319–326. Available at: https://doi.org/10.1038/s41586-024-08328-6.

Jumaah, M.A., Ali, Y.H. and Rashid, T.A. (2025) ‘Artificial Liver Classifier: A New Alternative to Conventional Machine Learning Models’. arXiv. Available at: https://doi.org/10.48550/arXiv.2501.08074.

Karar, M.E., El-Fishawy, N. and Radad, M. (2023) ‘Automated classification of urine biomarkers to diagnose pancreatic cancer using 1-D convolutional neural networks’, *Journal of Biological Engineering*, 17(1), p. 28. Available at: https://doi.org/10.1186/s13036-023-00340-0.

Laxminarayanamma, K., Krishnaiah, R.V. and Sammulal, P. (2022) ‘Enhanced CNN Model for Pancreatic Ductal Adenocarcinoma Classification Based on Proteomic Data’, *Ingénierie des systèmes d information*, 27(1), pp. 127–133. Available at: https://doi.org/10.18280/isi.270115.

LeCun, Y., Bengio, Y. and Hinton, G. (2015) ‘Deep learning’, *Nature*, 521(7553), pp. 436–444. Available at: https://doi.org/10.1038/nature14539.

Levine, S. *et al.* (2018) ‘Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection’, *The International Journal of Robotics Research*, 37(4–5), pp. 421–436. Available at: https://doi.org/10.1177/0278364917710318.

*Machine Learning vs. AI: Differences, Uses, and Benefits* (2025) *Coursera*. Available at: https://www.coursera.org/articles/machine-learning-vs-ai (Accessed: 11 March 2025).

Maftoun, M. *et al.* (2024) ‘Malicious URL Detection using optimized Hist Gradient Boosting Classifier based on grid search method’. arXiv. Available at: https://doi.org/10.48550/arXiv.2406.10286.

*Map of Computer Science* (2017). Available at: https://www.youtube.com/watch?v=SzJ46YA\_RaA (Accessed: 7 March 2025).

Ravikumar, C. *et al.* (2024) ‘A Comprehensive exploration of machine learning in early detection with a focus on lung and pancreatic cancer for revolutionizing cancer diagnostics’, in *2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*. *2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*, pp. 1–6. Available at: https://doi.org/10.1109/ICETCS61022.2024.10543936.

Russell, S.J. and Norvig, P. (2016) *Artificial intelligence : a modern approach*. Pearson. Available at: https://thuvienso.hoasen.edu.vn/handle/123456789/8967 (Accessed: 6 March 2025).

Sai, M.D.S. *et al.* (2024) ‘Early Detection: Machine Learning Techniques in Pancreatic Cancer Diagnosis’, *Journal of Engineering Research and Reports*, 26(5), pp. 175–182. Available at: https://doi.org/10.9734/jerr/2024/v26i51144.

Saxena, A. and Saxena, S. (2023) ‘Pancreatic Cancer Data Classification with Quantum Machine Learning’, *Journal of Quantum Computing*, 5(1), pp. 1–13. Available at: https://doi.org/10.32604/jqc.2023.044555.

Vaswani, A. *et al.* (2017) ‘Attention is All you Need’, in *Advances in Neural Information Processing Systems*. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html (Accessed: 6 March 2025).

*What Is Artificial Intelligence? Definition, Uses, and Types* (2024) *Coursera*. Available at: https://www.coursera.org/articles/what-is-artificial-intelligence (Accessed: 9 March 2025).