# **A Perceptron-based Approach for Diabetes Data Classification**

|  |
| --- |
| Aditya D Pandey  The University of Adelaide  Adelaide, South Australia, Australia  aditya.pandey@student.adelaide.edu.au |

**Abstract**

*The ABSTRACT is to be in fully-justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word “Abstract” as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. The abstract may be up to 3 inches (7.62 cm) long. Leave two blank lines after the Abstract, then begin the main text.*

Add here later

# Introduction

Diabetes is a long-term condition that impacts individuals across the globe (Khan et. al 2020). Precise categorization of diabetes information plays a role, in identifying it making diagnoses and devising effective treatment strategies (WHO 2019). Over the years the application of machine learning methods has demonstrated encouraging outcomes in analyzing medical data particularly when it comes to classifying diabetes (Hama Saeed 2023), (Patra & khuntia 2021), (Darolia & Chhillar 2022), (Rajni & Amandeep 2019) 14-45 from Ismail [19, 21, 29, 29, 40, 51, 57]. From Howlader. This study introduces an approach based on perceptron, for categorizing diabetes data by harnessing the capabilities of neural networks to enhance accuracy and efficiency.

# Method

In this section we will delve into the materials and methods utilized in the research study. Subsequently we will provide a breakdown of the dataset, methods and approach, in subsequent subsections.

## Dataset

The dataset utilized in this research corresponds to the Pima Indian community situated near Phoenix, Arizona, a dataset that has been continuously examined since 1965 by the National Institute of Diabetes and Digestive and Kidney Diseases (Smith et. al 1988)

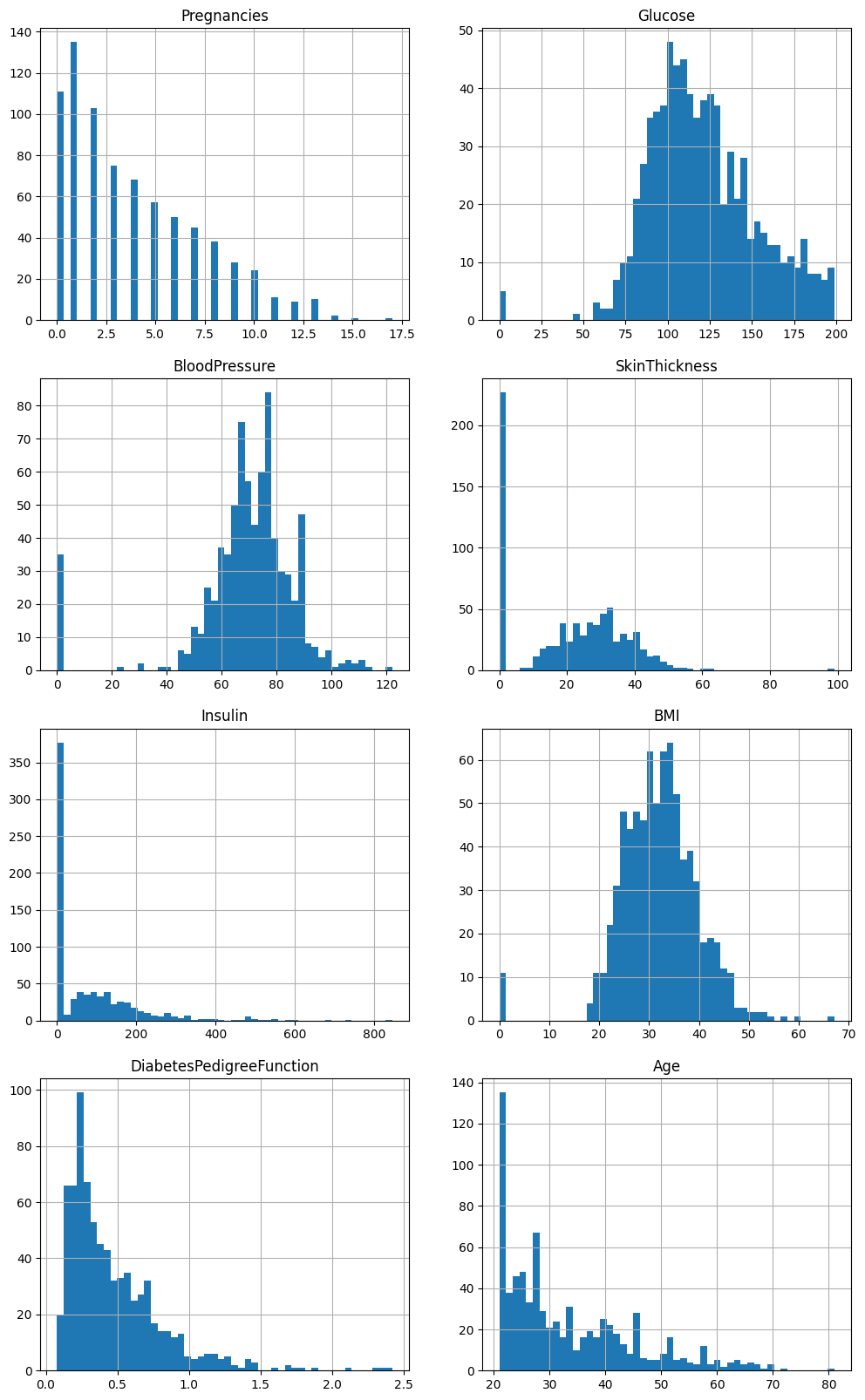


Figure 1: A histogram for each numerical attribute



Table 1: Description of all the PIDD attributes

PIMA Indian diabetes dataset (PIDD) consists of 9 attributes, 8 predictors and 1 class label. The dataset comprises of 768 women with age of more than 21 years. Above depicted are histogram plots and table with description and for each of attribute(s) of PIDD (Naz & Ahuja 2020 p. 395).

## Single-Layer Perceptron (SLP)

The Single-Layer-Perceptron is the most basic Artificial Neural Network architectures, the model was invented by Frank Rosenblat (Rosenblatt 1958). ~~A SLP can be used for classification on linearly separable problem.~~

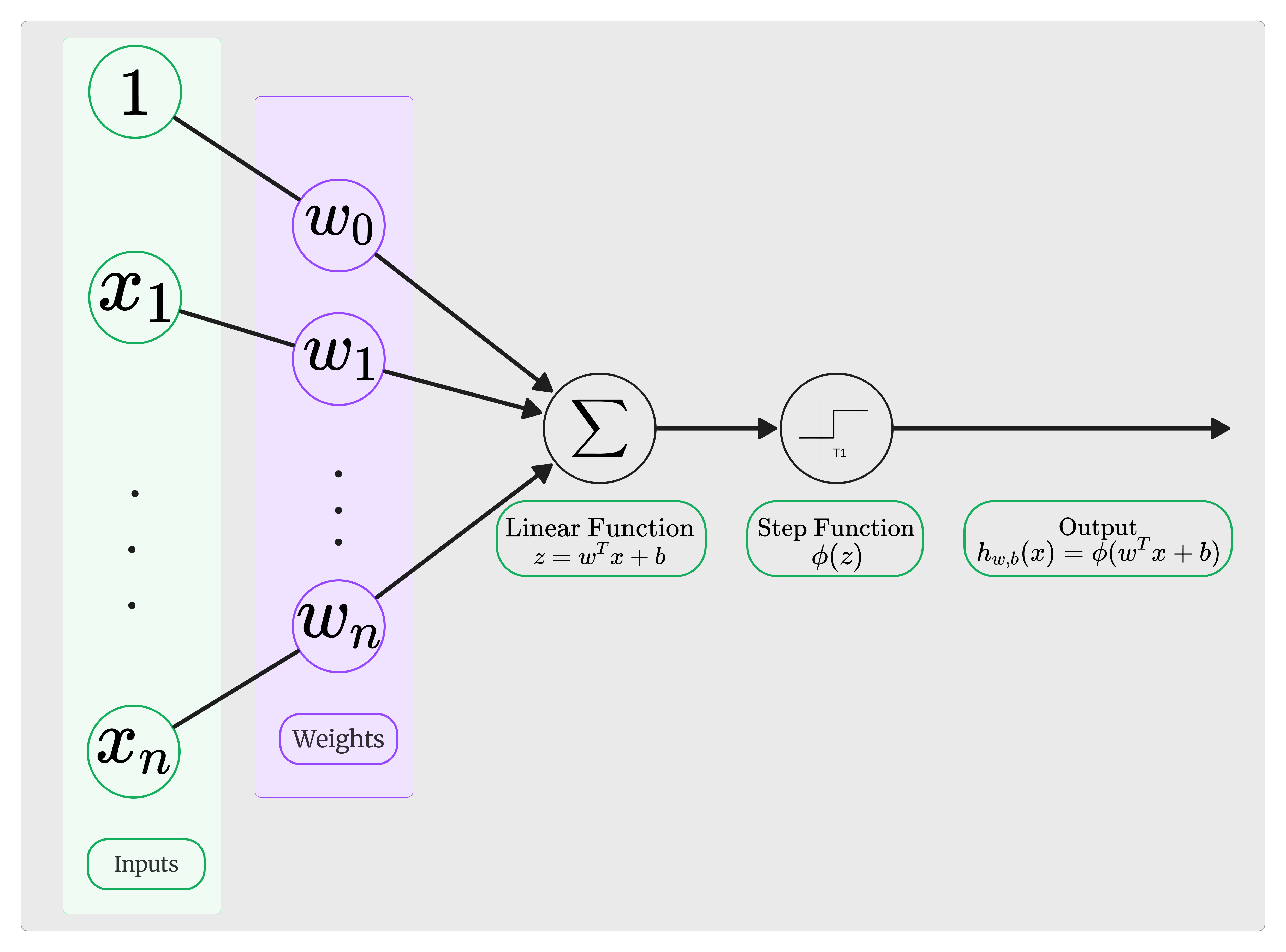
A Perceptron can be described as a linear threshold device that computes weighted sum of the coordinates of the pattern vector, compares the value with a threshold, and outputs +1 or -1 if the threshold is reached, threshold is identified as the activation function that we employ (Siu, Dembo & Kailath 1994 p.349). The below diagram depicts the concept.

Figure 2 : The Components of Single Layer Perceptron

SLP even though being the most elementary form of ANN still find use in multitude of fields of study, Alkhamees BF utilized optimized SLP for classification in context of fetal state detection (Alkhamees 2022, p. 244), SLP to analyze labroartory data (Forsström et. al 1995, pp.79-80), for prediction of Bankruptcy (Hu 2009, p. 1), and data security (Mualfah, Fatma & Ramadhan 2020, p. 3).

The core formula behind it is

As is depicted in the Figure 2. The inputs (x\_0, x\_1, . . . x\_n) are multiplied with the individual weights (1, w\_1, w\_2, . . .w\_n), further down the line they are summed up. Then the weighted sums are pass through a step function also known as activation function or transfer function, which as described below.

)

In order to simplify the calculation and reduce the processing delay, we employ the Sigmoid activation function a nonlinear activation function to provide clear probabilistic outputs, enabling straightforward interpretation for the binary classification.

3