1. **Introduction**

The objective of the network will be used to determine whether to take diabetes treatment to bring about more careful considerations of taking diabetes medicines which usually have many bad side-effects in the current diabetes field.

The potential user community is people who have family background diabetes history, old people with health concerns and people who generally worry about their health issues.

1. **Problem analysis**

The problem is whether to take diabetes treatments based on several observed conditions.

The potential factors are:

* Plasma glucose concentration a 2 hours in an oral glucose tolerance test
* Blood Pressure: Diastolic blood pressure (mm Hg)
* Insulin: 2-Hour serum insulin (mu U/ml)
* BMI: Body mass index (weight in kg/(height in m)^2)
* Diabetes Pedigree Function: Diabetes pedigree function
* Age: Age (years).

The dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective is to predict based on diagnostic measurements whether a patient has diabetes. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage. The dataset has been adopted for the use of this assignment.

The expert knowledge consulted are to know about background of diabetes. And also the criteria for changing numeric data to categorical variables based on some expert knowledge.

1. **Decision Network model**
2. **Model testing and evaluation**
3. **Conclusion**
4. **References**

Dataset resource

<https://www.kaggle.com/datasets/mathchi/diabetes-data-set>

glucose criteria (expert knowledge >140 is high)

<https://www.cdc.gov/diabetes/basics/getting-tested.html#:~:text=A%20fasting%20blood%20sugar%20level,higher%20indicates%20you%20have%20diabetes>.

Blood pressure (expert knowledge >80 is high)

<https://www.cdc.gov/bloodpressure/about.htm#:~:text=%2F80%20mmHg.%E2%80%9D-,What%20are%20normal%20blood%20pressure%20numbers%3F,less%20than%20120%2F80%20mmHg.&text=No%20matter%20your%20age%2C%20you,pressure%20in%20a%20healthy%20range>.

Insulin (avg of data excluding zeros, >156 is high)

Bmi (expert knowledge > 25 is high)

<https://www.calculator.net/bmi-calculator.html>

<https://www.escardio.org/The-ESC/Press-Office/Press-releases/Body-mass-index-is-a-more-powerful-risk-factor-for-diabetes-than-genetics>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4457375/#:~:text=The%20risk%20of%20developing%20DM,HR%3D2.51%20for%20women>).

DiabetesPedigreeFunction (likelihood so > 0.5 is high)

Age(general knowledge >50 is high)