

CS314:Big Data

Health and Big Data

Diabetes Analysis

Big Data | 28th November 27, 2017

|  |  |  |  |
| --- | --- | --- | --- |
| **SNo** | **Name** | **USN** | **Class/Section** |
| 1 | Bhavana NS Malepaty | 01FB15ECS068 | 5th Sem/Sec B |
| 2 | Bindu R | 01FB15ECS070 | 5th Sem/Sec B |

# Introduction

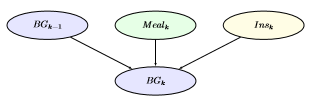
* The objective of the project is to develop a prediction system model for diabetic patient for their glucose levels
* The UCI Diabetes Repository, which had blood glucose data for 70 patients, was used. The attributes include blood glucose measured before and after a meal, insulin dosages at different times of the day, meal ingestion, and exercise activity

## Related work

IEEE Paper - A Probabilistic Framework for Blood Glucose Control in Diabetes

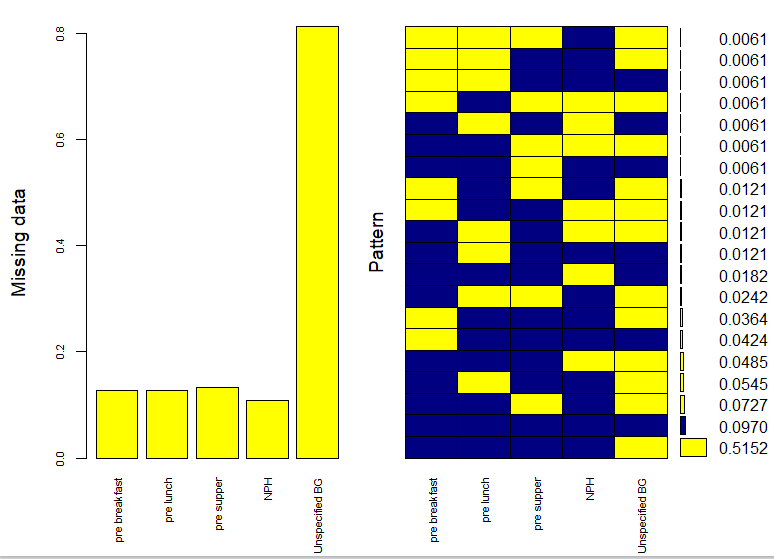
BG evolution should be described rather in terms of probabilistic models than in terms of deterministic ones. The paper explains a probabilistic model to ﬁnd an insulin dosing which offers the same quality of treatment than that of standard control strategies, but with less insulin used

Predicting the blood glucose through a probabilistic approach, using conditional probability, Bayesian Network, and Markov’s chain model



## ALGORITHM/DESIGN

The missing values are first imputed, using the MICE package



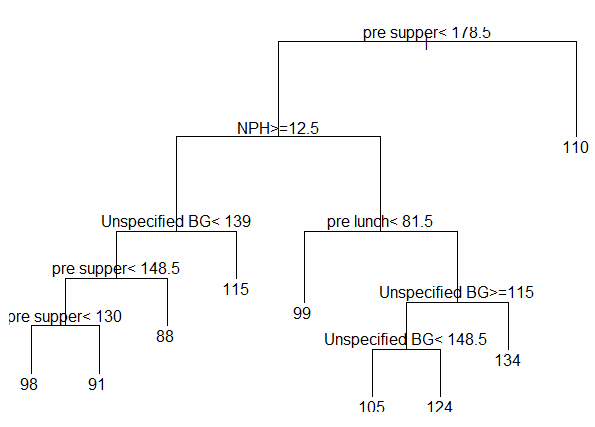
Two models were built

The dataset was split into training (90%) and testing (10%). The dependent variable chosen was pre-breakfast blood glucose

1. Linear Regression Model With SGD(Stochastic Gradient Descent)
2. Decision Tree Model

## EXPERIMENTAL RESULTS

1. Linear Regression Model
2. Decision Tree Model



## FUTURE ENHANCEMENTS

* Obtain non-intrusive glucose monitor
* Develop a multi-patient model, and introducing a new factor “insulin\_sensitivity”, that accounts for difference in response between different patients

## REFERENCES

* https://archive.ics.uci.edu/ml/datasets/diabetes