# Results and Discussion

The mRMR algorithm produces a sequential list of ten ranked features, listed in Table 2 that exhibit maximum dependency to the diabetic class. Notably, the highest ranking features involve the 2-hr values of the blood glucose and insulin measurements. The subject ethnicity is likewise a feature in predicting the future evolution of T2DM.

Table 2: List of ten most relevant features ranked by the mRMR algorithm

|  |  |
| --- | --- |
| Rank | Feature |
| 1 | AuC-Glu0-120 |
| 2 | ΔGlu120-0 |
| 3 | ΔGlu120-60 |
| 4 | ETHN |
| 5 | ΔIns120-0 |
| 6 | ΔGlu60-0 |
| 7 | ΔGlu30-0 |
| 8 | ΔGlu60-30 |
| 9 | ΔIns120-60 |
| 10 | ΔIns60-0 |

In all the classification experiments, we aimed to maximize the ability to correctly predict the diabetic class. The bar plots in Fig. 3 show the geometric mean (g-mean) of the sensitivity and specificity obtained from the linear and RBF kernels. For each number of features used, we selected the combination that generated the maximum g-mean. All the results presented here are averaged over 100 iterations of the respective classifiers. The accuracy and specificity of the best feature combinations are also shown in Fig. 3. A combination of four features, namely AuC-Glu0-120, ΔGlu120-0, ΔGlu120-60 and ΔGlu30-0 provided the best classification performance, with a g-mean of 0.89, accuracy of 0.9681, and sensitivity of 0.8045.

Table III presents a comparison of the generated SVM models with results obtained in other studies using the SAHS dataset. In addition to the SVM based prediction, we also explored the effect of balancing the class prevalence in the SAHS dataset prior to training. The majority class was randomly under-sampled, and 160 instances from each class were used. Although, a marginal increase in the classifier accuracy due to the introduction of an artificial bias was observed, the classifier sensitivity decreased which contraindicated the process of balancing the dataset.

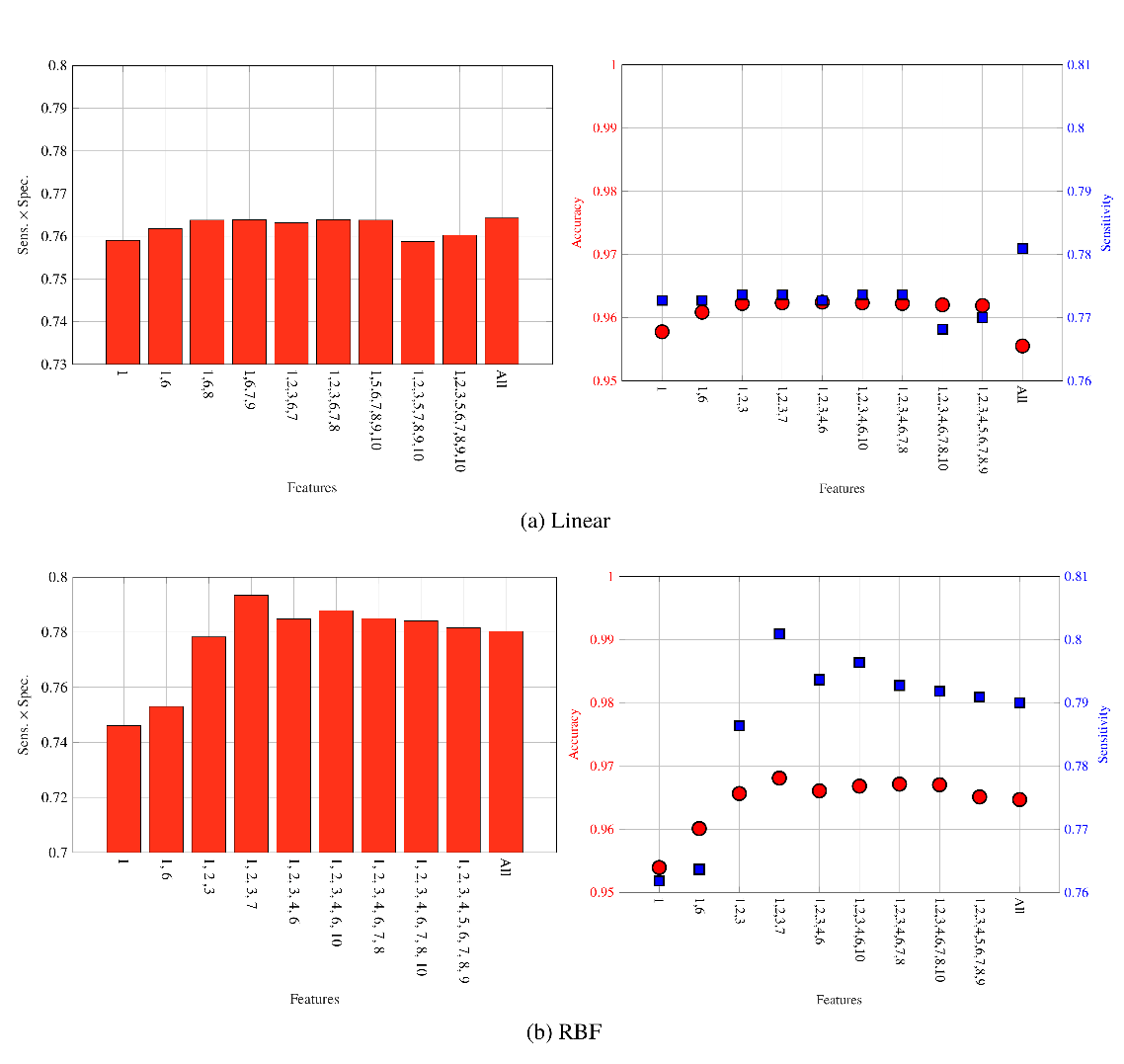


Figure 3: Geometric mean of sensitivity and specificity, and performance in terms of accuracy and sensitivity for the best feature combinations

We compared our results with the San Antonio diabetes prediction model (SADPM) [15], in which a person’s age, gender, ethnicity, fasting glucose level, family history, blood pressure, and cholesterol level were used to construct a logistic regression. It is notable that the SADPM has a very high sensitivity of 88.80%, however this increased prediction performance comes along with a very low accuracy of only 56.329%. In [16], a two-step approach was introduced that first used the SADPM risk score and then augmented it with the 1-hr blood glucose level. This strategy resulted in an improved accuracy but the sensitivity dropped to 77.70%

Table 3 also reveals that out of those subjects, diagnosed with prediabetes according to the WHO IGT criteria, only 33.93% developed diabetes between the baseline and the follow-up. Moreover, only 8.19% of the subjects were diagnosed with diabetes that matched one of the diagnostic criteria defined by the American Diabetes Association (ADA).

The non-linear SVM using the

Table 3: Validation performance classifiers

|  |  |  |  |
| --- | --- | --- | --- |
|  | Accuracy ± SD | Sensitivity ± SD | Specificity ± SD |
| Linear SVM (Balanced) | 97.18 % ± 1.48 % | 75.91 % ± 12.68 % | 100 % |
| Linear SVM (Unbalanced) | 96.20 % ± 1.94 % | 77.00 % ± 12.08 % | 98.75 % ± 1.35 % |
| SVM-RBF (Balanced) | 97.73 % ± 1.41 % | 80.00 % ± 12.03 % | 100 % |
| SVM-RBF (Unbalanced) | 96.81 % ± 1.40 % | 80.45 % ± 11.50 % | 100 % |
| SADPM [15] | 56.329 % | 88.80 % | 52.00 % |
| Two-step Approach [16] | - | 77.70 % | 77.40 % |
| IGT (PG120 > 140 & <200 mg/dL) |  | 33.93 % |  |
| ADA Criteria (PG120 > 200 mg/dL) |  |  |  |

# Conclusion

In this paper, we present a most-parsimonious set of features that are used to develop a non-linear SVM based future T2DM prediction model. The features were derived from the oral glucose tolerance test. Using a feature selection algorithm, we demonstrate that the features that are deduced from the blood glucose levels have are the strongest predictors of the future development of T2DM. Moreover, the performance of the presented prediction model is significantly better in terms of accuracy and sensitivity, as compared to other T2DM prediction schemes. In order to address the unbalanced nature of the SAHS dataset, we chose the geometric mean of sensitivity and specificity as the performance evaluation criteria. We also observed that balancing the class prevalence in the dataset did not result in performance improvement.

The principal contribution of this paper is that a T2DM prediction model based on the features derived only from the blood glucose levels measured during an OGTT significantly outperforms other models. The findings of this paper can provide a tool for the clinicians to screen individuals that are at an increased risk of developing T2DM in future.