**Interpretation and Discussion**

**EVALUATION OF MODELS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **SVM (Linear)** | **SVM (Polynomial)** | **SVM (RBF)** | **RFC** |
| **Accuracy** | 0.82 | 0.75 | 0.79 | 0.79 |
| **Precision** | 0.76 | 0.71 | 0.70 | 0.69 |
| **Recall** | 0.62 | 0.32 | 0.55 | 0.57 |
| **F1-Score** | 0.68 | 0.44 | 0.62 | 0.63 |
| **AUC-ROC** | 0.77 | 0.63 | 0.73 | 0.73 |
| **Mean accuracy** | 0.76 | 0.73 | 0.74 | 0.73 |
| **Standard Deviation** | 0.03 | 0.04 | 0.03 | 0.05 |

Based on our evaluation of the SVM models, it appears that the SVM with the Linear kernel performed the best, achieving an accuracy of 82% on the test set, a cross-validated mean accuracy of 76% and standard deviation of 0.03. The RBF kernel also performed reasonably well, achieving an accuracy of 79% on the test set, a cross-validated mean accuracy of 74% and standard deviation of 0.03. The polynomial kernel performed the worst, with an accuracy of 75% on the test set, a cross-validated mean accuracy of 73% and standard deviation of 0.04.

Based on our evaluation of the Random Forest models, it appears an accuracy of 79% on the test set, a cross-validated mean accuracy of 73% and standard deviation of 0.05.

It is worth noting that the performance of the models could potentially be improved by further tuning the hyperparameters or by using other machine learning algorithms.

In terms of healthcare interventions for the Pima Indian population, the best-performing model (i.e., SVM with the linear kernel) could be used to identify individuals who are at high risk of developing diabetes based on their medical history and demographic information. Healthcare providers could then take targeted preventive measures, such as providing dietary and lifestyle recommendations, to help these individuals reduce their risk of developing diabetes. Additionally, the model could be used to identify factors that are most strongly associated with diabetes risk in this population, which could inform public health interventions and policies aimed at reducing the prevalence of diabetes in the Pima Indian population.