Chapter 12
Machine Learning Through Early Diabetes Detection: Evaluating Random Forest Classifier Performance

Omkar Shridhar Bholankar
https://orcid.org/0009-0000-3581-0124
Ajeenkya D.Y. Patil School of Engineering, India

Shreya Bhalerao
https://orcid.org/0009-0009-3645-5764
Ajeenkya D.Y. Patil School of Engineering, India

Shravani Bhalerao
https://orcid.org/0009-0009-3645-5764
Ajeenkya D.Y. Patil School of Engineering, India

Sweta Anand Wankhade
https://orcid.org/0009-0006-3667-5781
Ajeenkya D.Y. Patil School of Engineering, India

ABSTRACT

Diabetes is among the life-threatening diseases affecting millions of people worldwide. Detection done early and accurate prediction of diabetes can significantly improve patient outcomes by enabling timely intervention and lifestyle modifications. In recent years, Machine Learning (ML) techniques have been extensively used in medical diagnostics due to their ability to analyze complex patterns in healthcare data. This study focuses on evaluating the performance of the Random Forest (RF)

DOI: 10.4018/979-8-3373-4094-4.ch012

algorithm. The dataset including feature scaling, handling missing values and splitting into training and testing sets to ensure optimal mode performance. The results are evaluated based on performance metrics like Accuracy, Recall, Precision, and F-Measure that are derived from the confusion matrix. The experimental results proved that the best accuracy goes for Random Forest (RF).

1 INTRODUCTION

The global rise of diabetes mellitus represents one of the most significant and relentless public health challenges of the 21st century. It is a life-threatening chronic disease that affects millions of people worldwide. According to the International Diabetes Federation, an estimated 537 million adults were living with diabetes in 2021, a number projected to soar to 783 million by 2045. This silent epidemic is not just a health crisis but also a staggering economic one, with global health expenditure on diabetes reaching an estimated USD 966 billion. This figure fails to capture the immense indirect costs associated with lost productivity, disability, and the profound impact on national economies. Characterized by high blood glucose levels resulting from insufficient insulin production or ineffective use, diabetes, if not managed with constant vigilance, can lead to a cascade of severe and debilitating health complications. These include cardiovascular diseases, nerve damage (neuropathy), kidney failure, and vision impairment leading to blindness. As the global incidence of this disease continues its upward trajectory, the locus of care is undergoing a seismic shift, moving from the structured, professional environment of the hospital to the complex and often-unsupported setting of the home. This transition places an unprecedented level of responsibility onto the shoulders of patients and, most critically, their families.

Within this paradigm of home-based care, family caregivers emerge as the linchpins of chronic disease management. These unsung heroes—spouses, adult children, and other relatives—form the invisible frontline of our healthcare system, providing essential daily support that is both complex and demanding. The burden they carry is multifaceted, extending far beyond simple companionship. The physical burden is relentless, involving a daily regimen of tasks such as monitoring blood glucose levels, preparing carefully balanced meals, ensuring medication is taken correctly and on time, and managing the physical symptoms of the disease. This can lead to chronic fatigue, sleep deprivation, and a decline in the caregiver's own physical health.

Consider the hypothetical but deeply representative story of Anjali, a working professional in her forties who cares for her aging father, Mr. Sharma, recently diagnosed with Type 2 diabetes. Her day begins before dawn, preparing a low-glycemic breakfast for her father and checking his morning blood sugar. She meticulously