**Performance Assessment**

OFM4 — OFM4 Task 2: Dimensionality Reduction Methods

Bader Ale

Department of Information Technology, Western Governors University

D212 Data Mining II

2024

Table of Contents

[Part I: Research Question 3](#_Toc179451279)

[Part II: Technique Justification 3](#_Toc179451280)

[Part III: Data Preparation 5](#_Toc179451281)

[Part IV: Analysis 7](#_Toc179451282)

[Part V: Data Summary and Implications 7](#_Toc179451283)

# Part I: Research Question

# The purpose of this report is to explore the following research question: *"What factors in a hospital's medical dataset explain the biggest differences in patient readmission outcomes?"* By using Principal Component Analysis (PCA), the goal is to identify the most important variables that influence differences in readmission rates among patients. Understanding these key factors can help hospital staff make informed decisions to improve patient care and lower readmission rates. This question is relevant to real-world healthcare settings, where efficient resource management and focused interventions are essential for better patient outcomes and smoother hospital operations.

The goal of this analysis is to simplify the hospital's medical dataset while retaining the most important information. Specifically, PCA will be used to reduce the original set of variables to a smaller set of principal components that capture the majority of the variance in the data. The aim is to identify the components that have the most influence on patient readmissions, helping healthcare providers focus on these factors for targeted interventions. This approach helps make the dataset more manageable, allowing for more effective analysis and actionable insights to reduce readmission rates.

# Part II: Method Justification

PCA analyzes the dataset by transforming the original features into a new set of uncorrelated variables, called principal components, which capture the maximum variance in the data. By doing this, PCA reduces the dimensionality of the data while retaining as much information as possible. This process allows us to simplify complex datasets and focus on the components that contribute most significantly to differences in patient readmission rates.

The expected outcome of using PCA is a smaller number of components that still explain the majority of variance in the original dataset. This helps identify key factors driving patient readmission, making it easier for hospital administrators to prioritize areas for improvement and allocate resources effectively (Fonseca, 2023).

# Part III: Data Preparation

# Part IV: Analysis

Works Cited

Fonseca, M. (2023, October 19). *editage insights*. Retrieved November 2024, from An introduction to Principal Components Analysis for biomedical researchers: https://www.editage.com/insights/an-introduction-to-principal-components-analysis-for-biomedical-researchers