



University of Zabol
Faculty of new technology

The Thesis Submitted for M.Sc. Degree of Field of bioinformatics

Title:

**Investigating the performance of k-Nearest Neighbor classifier based on two
new distance functions in diagnosis of different cancers**

Supervisor:

Dr. Mohammad Reza Galavi

Dr. Rezvan Ehsani

Adviser:

Dr. Mansour Ghaffari Moghaddam

By:

Mohsen Mollashahi

March, 2021

Abstract

One of the simplest and yet comparable efficiency and even in some optimal cases, is the complex and modern algorithms available in machine learning domain and especially the classification problem, the k - nearest neighbor (KNN) algorithm. The core of this algorithm is directly dependent on the measurement of distance or similarity between the experimental sample and the training samples. Due to the frequency and variety of distance functions and topology and spatial geometry in which the data are located, especially in high dimensions, determining the optimal metric and testing new metrics is necessary to achieve maximum efficiency. In this study, the aim is to use the two new distance functions of Sobolev and Fisher for the first time in the KNN algorithm and to evaluate and compare their performance with existing distance functions for data, especially data related to the field of cancer diagnosis.

Keywords: K-Nearest Neighbor, Classification, Distance Fisher, Sobolev, Cancer