# ABSTRACT

Abstract of the thesis should be in one page with the following parameters:

|  |  |
| --- | --- |
| **For Heading:**  **Font Size:** 20  **Font Name:** Times New Roman  **Alignment:** Center | **For Text:**  **Font Size:** 12  **Font Name:** Times New Roman  **Alignment:** FullJustified |

**Sample Abstract:**

# ABSTRACT

**[Font Size 20]**

Modeling and finding of repetitive patterns, called motifs in unaligned DNA sequence is a fundamental problem in computational biology with important applications in understanding gene regulation. These sets of similar subsequences tend to indicate regions that have the same or similar, biological function and therefore hint at the purpose and structure of uncharacterized sequence.

Biological approaches for motif finding problem are tedious and time-consuming. So, finding motif on such large-scale need to be efficient and also be sensitive enough to recover subtle but significant motif.

In this work, the performance of a motif finding algorithm, SPSTAR is investigated. This algorithm is implemented to solve planted motif problem and its performance is tested on synthetic data. A modified version of the SPSTAR algorithm is implemented and its performance is tested in the same setting. It is found that the performance of the modified SPSTAR algorithm is better than the existing SPSTAR algorithm in terms of time complexity. A new algorithm for finding motif using Suffix Tree is also proposed. This algorithm can find best motif available on a given set of input sequences up to a certain substitution limit. It also requires less computational time in motif finding than SPSTAR. Due to the characterization of the planted motif, applications of these motif-finding algorithms in biological domain are limited.

**[Font Size 12]**