Title Page

Design & Implementation of Clustering techniques based Simulator

Project Report submitted in partial fulfillment of the requirement for the degree of

Bachelor of Technology.

in

Computer Science & Engineering

under the Supervision of

Dr. Pardeep Kumar

By

Prateek Tiwari (091265)

to

Jaypee University of Information and Technology
Waknaghat, Solan – 173234, Himachal Pradesh

CERTIFICATE

This is to certify that project report entitled **" Design & Implementation of Clustering techniques based Simulator",** submitted by **Prateek Tiwari** in partial fulfillment for the award of degree of Bachelor of Technology in Computer Science & Engineering to Jaypee University of Information Technology, Waknaghat, Solan has been carried out under my supervision.

This work has not been submitted partially or fully to any other University or Institute for the award of this or any other degree or diploma.

Date: 12-05-2013

Supervisor's Name

Dr. Pardeep Kumar

Designation

Assistant Professor

ACKNOWLEDGEMENT

I would like to extend my sincere gratitude to our project mentor Dr. Paedeep Kumar, Who provided us inspiring guidance and gave us an insight into the latest technical Development going in the vast field learing. I would also like him to thank for devoting considerable part his precious time to give us a better knowledge of the field data mining, without his support my project would not have been possible.

We would like to thak JUIT , Department of Computer Science & Engineering and all other members who provided me this golden opportunity and supported me to take another great step towards the completion final year project. I owe our heartiest thanks to Brig. (Retd.) S.P. Ghrera (H.O.D CSE/I.T Department) who always inspired confidence in me to take initiative. He has always been motivating and encouragaing.

Date: **12-05-2013** Name of the student

Prateek Tiwari

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

Data mining and Knowledge discovery has several important application areas. Data mining and knowledge discovery have been topics considered at many AI, database and statistical conferences. Knowledge discovery generally refers to the process of identifying valid, novel and understandable patterns. Knowledge discovery from large databases, often called data mining, refers to the application of the discovery process on large databases or datasets. The discovery process can be broken into several steps, including: developing an understanding of the application domain; creating a target data set; data cleaning and preprocessing; finding useful features with which to represent the data; data mining to search for patterns of interest; and interpreting and consolidating discovered patterns. Data mining and knowledge discovery in databases have been attracting a significant amount of research, industry, and media attention of late. What is all the excitement about? This article provides an overview of this emerging field, clarifying how data mining and knowledge discovery in databases are related both to each other and to related fields, such as machine learning, statistics, and databases. The article mentions particular real-world applications, specific data-mining techniques, challenges involved in real-world applications of knowledge discovery, and current and future research directions in the field.

Data mining on large databases has been a major concern in research community, due to the di_culty of analyzing huge volumes of data using only traditional OLAP tools. This sort of process implies a lot of computational power, memory and disk I/O, which can only be provided by parallel computers. We present a discussion of how database technology can be integrated to data mining techniques. Finally, we also point out several advantages of addressing data consuming activities through a tight integration of a parallel database server and data mining techniques