**Principles of Green Data Mining**

**ABSTRACT**

This paper develops a set of principles for green data mining, related to the key stages of business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The principles are grounded in a review of the Cross Industry Standard Process for Data mining (CRISP-DM) model and relevant literature on data mining methods and Green IT. We describe how data scientists can contribute to designing environmentally friendly data mining processes, for instance, by using green energy, choosing between make-or-buy, exploiting approaches to data reduction based on business understanding or pure statistics, or choosing energy friendly models.

**Existing System**

Data scientists often leverage a large pool of computational resources using sophisticated and computationally costly machine learning techniques to extract knowledge and insights from data. Though existing processes such as the Cross Industry Standard Process for Data mining (CRISP-DM) [61] provide some guidance on how to execute a data mining project, the skills of a data scientist heavily rely on creativity [53], involving many degrees of freedom, often including the choice of tools, models, and data sources

**Disadvantages**

1. there is no clear consensus about which model is best for a task
2. This is supported by the “no free lunch” theorem stating that any algorithm outperforms any other algorithm on some datasets

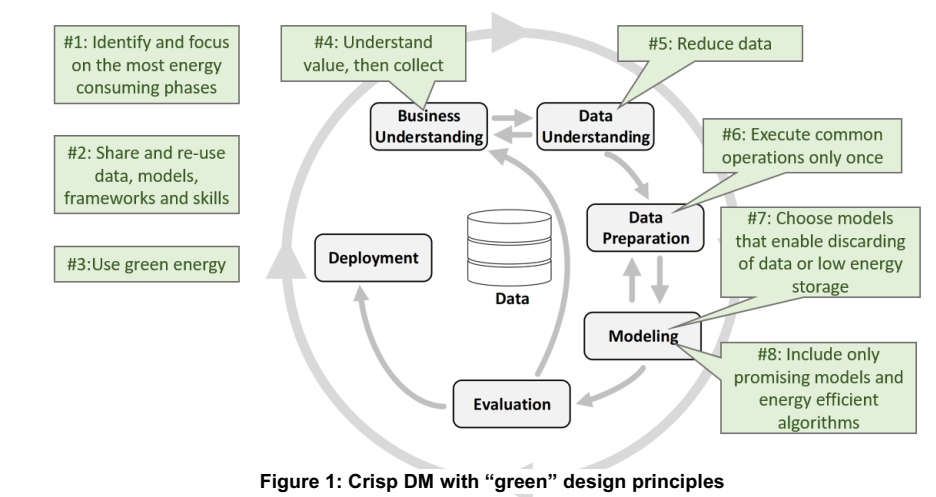
**Proposed System**

In this paper, we develop guidelines for data scientists to implement more environmentally friendly practices that can complement technology-focused perspectives aiming to design more energy efficient IT-based systems. Specifically, we are focusing attention on one important area of data science—data mining. Data mining can be described as knowledge discovery from data or in terms of different activities as collecting, cleaning, processing, analyzing and gaining useful insights from data. We ask: How can data scientists implement more environmentally friendly data mining processes? The remainder of this paper is structured as follows. We first describe our methodology. We then review the data mining process and develop a set of principles for green data mining. We conclude by discussing limitations and future work

**Advantages**

1. More environmentally friendly practices that can complement technology-focused perspectives aiming to design more energy efficient IT-based systems.
2. cleaning, processing, analyzing and gaining useful insights from data

**System Architecture**

****

# Hardware Requirements:

# Processor - Pentium –IV

* Speed - 1.1 GHz
* Ram - 256 MB
* Hard Disk - 20 GB
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

**Software Requirements:**

* Operating System - Windows XP
* Coding Language - java