Chapter 1
Optimizing eCommerce Data: Effective Approaches for Data Collection, Cleansing, and Preprocessing

Chirag Sharma
Chitkara Business School, Chitkara University, Punjab, India

Amanpreet Kaur
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Priyanka Datta
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Yonis Gulzar
https://orcid.org/0000-0002-6515-1569
College of Business Administration, King Faisal University, Al Ahsa, Saudi Arabia

ABSTRACT

The process of gathering and analyzing data, on variables in a manner to address research inquiries, test hypotheses and evaluate results is referred to as data collection. After obtaining the required data the next important step is pre-processing it. Data pre-processing involves converting data into a dataset before running algorithms. It prepares the dataset by handling missing values, noisy data and inconsistencies. Sometimes additional adjustments such as handling outliers and

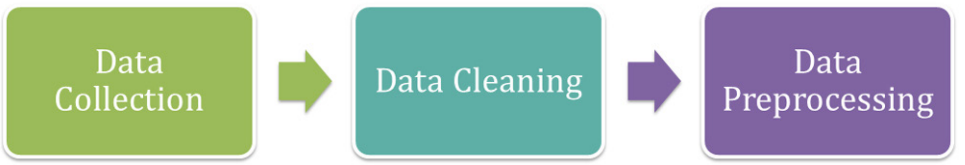
DOI: 10.4018/979-8-3693-5718-7.ch001

creating subcategories are needed for generating outcomes. This process can be seen as a step in organizing data effectively. The methods mentioned above demonstrate that research entails steps typically carried out manually or using applications by researchers. This article will explore an instances where Artificial Intelligence (AI) and Machine Learning (ML) are utilized for cleaning and pre-processing data. By examining the data AI and ML can detect patterns efficiently. Visual representations play a role, in understanding customer satisfaction and happiness levels within a shopping platform

1. INTRODUCTION

In today's world driven by data, e commerce companies thrive on the information obtained from customer behaviour, product trends and market dynamics. However, the data collected from sources is often messy, incomplete and inconsistent. To fully utilize this data's potential businesses, need to undergo a process. Data preparation. This process involves three steps; gathering data cleaning it up and pre-processing it as illustrated in Figure 1. Each step is essential, in converting data into insights that can guide decision making and boost e commerce success.

Figure 1. Phases of converting raw data into actionable insights



1.1 Data Collection

Data collection forms the first step in the data preparation process (Taherdoost, 2021). Gathering information, from sources is important for understanding the goals of a research project, as shown in Figure 2. Taking the time to plan and decide on the type of data needed is crucial as it shapes the direction of the research. This initial