CHAPTER I

INTRODUCTION

Over recent years, technological disruption has introduced innovative credit scoring methods that are revolutionizing how financial institutions assess customer creditworthiness. Innovative credit scoring involves using artificial intelligence, machine learning, and alternative data sources (utility payments, mobile phone bill settlements, social media activity, online activity, and e-commerce purchase history) to expand credit access for individuals without traditional credit histories (Wijaya, 2023). Lenders can use machine learning technology and advanced AI algorithms to reduce expenses associated with credit decisions and decrease operational costs by leveraging these non-traditional data sources (Jagtiani & Lemieux, 2019).

Generation Z may lack extensive credit histories but are digitally savvy (IDN Research Institute, 2023). Traditionally, financial institutions evaluate borrowers' creditworthiness based on their credit history, demographic data, and behavior (Niu et al., 2019). However, this traditional approach may exclude some individuals who do not have a credit history, especially affecting Generation Z. Despite being digitally savvy, they may still face challenges in the standard credit system. Recognizing the potential of unconventional data sources is crucial for bridging the gap and expanding access to credit. Data is the backbone of innovative credit scoring (Wijaya, 2023). Financial institutions can use social media data to increase credit access for borrowers who do not have traditional credit histories. This approach not only addresses the need for greater financial inclusion but also aligns with the characteristics of Generation Z.

Social media as a tool to assess credibility is valuable due to its ability to represent user behavior (Ramadhani et al., n.d.). LinkedIn is a professional social media platform that offers valuable insights into individuals and their networks (Bradbury, 2011; Puteri Ramadhani et al., n.d.). LinkedIn is suitable for describing social network data to determine creditworthiness by analyzing a borrower's connections (Niranjan et al., n.d.). In addition, demographic data such as age, education, salary, and gender can also be extracted from LinkedIn profiles (Bradbury, 2011; Guo et al., 2016). Moreover,

we can gain insights into an individual's personality by analyzing user-generated content such as textual data on LinkedIn profiles (Alamsyah et al., 2018; Ramadhani et al., n.d.).

Financial institutions employ the 5C analysis principle to measure a borrower's capacity and intention to repay borrowed funds to mitigate the risk of non-performing loans when the borrower does not meet the necessary criteria (Rian Basori et al., n.d.). The 5C analysis measures an individual's character, capacity, capital, collateral, and condition to determine their creditworthiness. Hence, this study will apply three of the five principles to evaluate borrowers' creditworthiness by leveraging data from social media. The character will analyze user-generated content to assess the borrower's personality (Ramadhani et al., n.d.). Afterward, we consider individual conditions by analyzing demographic data and evaluating their capacity within social networks (Orlova, 2021; Puteri Ramadhani et al., n.d.).

Employing social media data for credit scoring presents challenges as the absence of social media activity can bias assessments of individual creditworthiness. The lack of online presence may misrepresent an individual's financial behavior and risk. Exploring classification methods and careful data preprocessing is necessary to mitigate bias and ensure fair and accurate credit evaluations. In this study, we find it crucial to assess the feasibility of using classification methods for applying a predictive analysis model (Anil Kumar et al., 2022; Ledhem, 2022).

We can achieve personality measurement through interviews, questionnaires, and integrity tests (*Handbook of Employee Selection*, n.d.). These methods provide different approaches to evaluate personality traits and assess individuals in various contexts. Previous studies have connected creditworthiness to personality traits and demographic factors when traditional credit history is unavailable (Ramadhani et al., n.d.). However, they have not explored the social network features to create a comprehensive creditworthiness model. We are adding social network features that fulfill the 5C's principle of credit with a particular focus on the 'Capacity' to strengthen social media data as an alternative source in cases where there is a lack of credit history. The Big Five Model is a widely recognized personality trait assessment [Roberts et al.,

2023. The Big Five model evaluates characteristics of extraversion, agreeableness, conscientiousness, neuroticism, and openness. To measure personality using this model, we utilized a Personality Measurement Platform called Platform Pengukur Kepribadian, the Indonesian language version of the platform (Alamsyah et al., 2021).

We propose a credit scoring model that utilizes social media data from LinkedIn with an ensemble approach to enhance performance and determine a person's creditworthiness. Demographic, personality traits, and social network data from LinkedIn are utilized. Our study primarily concentrates on the performance and characteristics of learning models and algorithmic methods. We assess the performance of the mentioned algorithm using accuracy, precision, recall, F1-score, and ROC/AUC values as evaluation measures. The ensemble approach, which combines all base classifiers is used to compare the base classifiers individually.