Chapter 1

Introduction

1.1 Background

Kidneys are organs responsible for the body's waste and toxin removal, making them a crucial player in the excretory system (Gounden et al., 2024). A dysfunction and structural anomaly occurring in the kidneys will lead to various kidney diseases (KDs) according to the KDIGO (Kidney Disease: Improving Global Outcomes) guidelines. Kidney diseases can be classified according to the cause, severity of outcome, structural abnormality, and duration (Lameire et al., 2021). Notably, in terms of duration, KDs can be classified into two classes: acute kidney disease (AKD) and chronic kidney disease (CKD) (Levey et al., 2013).

Commonly overlooked as a coexisting condition with diabetes or hypertension, kidney diseases (KDs) are complex diseases that contribute to global morbidity and mortality. KDs elevate the risks associated with 5 major lethal diseases: cardiovascular diseases (CVD), malaria, diabetes, HIV (Human Immunodeficiency Virus), and hypertension (Luyckx et al., 2018). Reported by the Global Burden of Disease (GBD) study in 2015, it was found that death caused by kidney failure increased by 32% over 10 years, affecting up to 1.2 million people (Wang et al., 2016). Furthermore, kidney diseases are predicted to continue to be a global health concern in 2050. According to a recent GBD study, it was predicted that chronic kidney disease would significantly rise in ranking as a leading cause of death by 2050 (Vollset et al., 2024). Additionally, research notes that about 850 million people suffer from kidney diseases globally, especially those in LICs (low-income countries) and LMICs (lower-middle-income countries), due to the low availability of KD diagnosis and treatment in their respective countries (Francis et al., 2024). As the knowledge of its expected burden increases, it becomes apparent that the initial identification of kidney dysfunction is crucial.

Kidney function is typically assessed by measuring the patient's creatinine level, whether through blood tests or urine sampling. The normal range for creatinine levels is highly dependent on the age and sex of the patient. An excessively low or high creatinine level can indicate kidney dysfunction, and a high creatinine level suggests GFR (glomerular filtration rate) decrease, which is a standard marker for CKD (Shahbaz et al., 2024). Aside from blood tests and urinalysis, imaging techniques and biopsies are occasionally used for detailed diagnosis of KDs (Chouhan et al., 2023).

However, these diagnostic methods, particularly biopsies, can lead to some medical concerns, such as pain, bleeding, infections, and other minor complications due to technical flaws (Kajawo et al., 2021; Wu et al., 2025; Yahata et al., 2022). Aside from medical concerns, these clinical tests require higher labor, financial costs, and longer duration while receiving either outpatient or inpatient treatments (Bedetti et al., 2008; Jørgensen & Lind, 2022; Milani & Jialal, 2023). Consequently, due to these disadvantages, there is an increasing demand for research on effective and highly accessible non-invasive diagnostic tests, like an electrocardiogram (ECG). The simplicity, time efficiency, and cost efficiency of ECGs are a great asset to expand their diagnostic abilities for other organ dysfunctions or diseases aside from CVDs (Mamun & Elfouly, 2023). Recent studies have shown that the implementation of artificial intelligence has expanded the application of ECG in medical diagnostics for both cardiac and non-cardiac diseases, denoting the growing potential of AI usage in medical diagnostics (Martínez-Sellés & Marina-Breysse, 2023; Sau et al., 2024).

However, regardless of the potential, there are some major concerns with AI-aided medical diagnosis. One of them is caused by imbalanced medical data between classes for model predictions. Such imbalanced data will result in biased and suboptimal AI model performance, reducing the credibility of the generated prediction model (Cross et al., 2024). In medical settings, the two classes are usually the healthy (normal) and unhealthy (abnormal) patients, and the healthy samples are usually considered the majority class. Hence, the algorithms would naturally prioritize this group, leading to the neglect of the minority unhealthy sample, which is supposed to be the main focus (Salmi et al., 2024). Therefore, there is a necessity to overcome this common problem to allow the full potential of AI usage in medical diagnostics, allowing effective diagnosis even for seemingly non-related biological systems.

1.2 Objective

This research aims to tackle the common challenge of imbalanced data in medical settings, particularly computer-aided diagnosis. Additionally, this research also aims to reveal the potential of utilizing electrocardiogram data to assess kidney function using artificial intelligence (AI). Overall, the objective of this research is to explore the usage of AI in developing prospective medical diagnostics models that would be more efficient in cost, time, and labor.

1.3 Hypothesis

Considering both the background and objectives of this research, it centrally focuses on how to overcome imbalanced data in computer-aided medical diagnostics. Thus, this research hypothesizes that downsampling or balancing the data to make it balanced will provide better results compared to utilizing imbalanced data in artificial intelligence (AI) models. Additionally, it is also hypothesized that electrocardiogram (ECG) patterns will be able to identify abnormalities in kidney function.