A PREDICTIVE CLASSIFICATION MODEL FOR RUNNING INJURY

By:
DEVESH RAJ A/L GANESAN
(Matric number: 144432)

Supervisor:
ASSOCIATE PROFESSOR DR LOH WEI PING

This dissertation is submitted to

Universiti Sains Malaysia

As partial fulfilment of the requirement to graduate with honours degree in

BACHELOR OF ENGINEERING (MECHANICAL ENGINEERING)

School of Mechanical Engineering

Engineering Campus

Universiti Sains Malaysia

EMD452 : FINAL YEAR PROJECT, 2021/2022

ENDORSEMENT OF CORRECTION FORM

Section 1 (To be completed by Student)

Full name: Devesh Raj a/l Ganesan

Matric No: 144432 Degree Program: Mechanical Engineering

Date and Venue of Viva: 1st August 2022 and Online Viva through Google Meet

Thesis's Title: A predictive classification model for running injury

____

____

____

2nd August 2022

(Student's Signature) (Date)

Section 2 (Verification by Supervisor)

I AP Dr. Loh Wei Ping, hereby declare that I am satisfied with the final version of the FYP report made by the student and therefore agree for the student to submit his/her report.

PROF. MADYA DR. LOH WEI PING
Pensyarah
Pusat Pengajian Kejuruteraan Mekanik
Kampus Kejuruteraan
Universiti Sains Malaysia
14300 Nibong Tebal, Pulau Pinang.
Tel: 604-599 6397 / 604-599 6394
Fax:604-599 6912
Emel: meloh@usm.my

5/8/2022

(Signature and Stamp) (Date)

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed………………………………………………(DEVESH RAJ A/L GANESAN)

Date………………………………………………………………….…...(25/7/2022)

Statement 1:

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by giving explicit references. Bibliography/references are appended.

Signed………………………………………………… (DEVESH RAJ A/L GANESAN)

Date…………………………………………………………………………...(25/7/2022)

Statement 2:

I hereby give consent for my thesis, if accepted, to be available for photocopying and for interlibrary loan, and for the title and summary to be made available outside organizations.

Signed………………………………………………… (DEVESH RAJ A/L GANESAN)

Date…………………………………………………………………………...(25/7/2022)

ACKNOWLEDGEMENT

Firstly, I would like to thank the Universiti Sains Malaysia for providing me with a conducive learning environment and technical assistance throughout my Final Year Project study. I would also like to express my sincere gratitude to my supervisor, Associate Professor Dr Loh Wei Ping, for her guidance throughout this project. From the initial data collection stage towards achieving the objectives of my research study, she gave me inspiration and endless motivation. I grasp better knowledge on the research methodologies from the data mining courses she taught and occasional conference calls. Besides, she also taught me the research knowledge and skills that are useful for the project completion. This is definitely useful and applicable for my future career as an engineer. Furthermore, this project itself also provided me the awareness regarding running injuries and its severity. Finally, I would like to thank my friends, Amy Syakilla binti Muhammad Adnan Adanan Suppiah, Illayakantthan a/l Rajogoval, Thivines a/l Murugaya and Yashwini A/P Thamil Selvam who helped in my research study progress with their continuous support and faith in me throughout the highs and lows of my study life.

ABSTRAK

Kecederaan yang berkaitan dengan larian (adalah sakit muskuloskeletal pada bahagian bawah anggota badan yang menyebabkan sekatan atau pemberhentian larian. Kecederaan larian keseluruhannya telah dikaji daripada segi faktor penyebab serta kesilapan latihan. Pelbagai model telah direka untuk menangani isu ini, namun peratusan kecederaan masih membimbangkan. Oleh itu, matlamat kajian ini adalah untuk menentukan penentu kecederaan larian, mengklasifikasikan data larian mengikut tahap keterukan dan untuk membangunkan model pencegahan kecederaan larian yang terdiri daripada tahap keletihan/kecederaan. Dua kajian kes yang berkaitan dengan kecederaan larian telah diperoleh daripada domain umum. Pendekatan perlombongan data telah dijalankan untuk pra-pemproses data dan mengklasifikasi data kepada tiga tahap kecederaan: risiko rendah, sederhana, dan tinggi dengan bantuan perisian Waikato Environment for Knowledge Analysis (WEKA) versi 3.8. Algoritma-algoritma J48, SMO, Random Forest, dan Simple Logistic telah digunakan bagi mod pengesahan bersilang 10 ganda dengan menggunakan algoritma ZeroR sebagai penanda aras. Hasil kajian memaparkan bahawa ketepatan pengelasan yang diperoleh adalah dalam julat 70% hingga 100%.

ABSTRACT

Running- related injury is musculoskeletal pain in the lower limbs that causes a restriction on or stoppage of running. Running injuries have been collectively studied in terms of the attributing factors as well as faulty trainings. Various models have been devised to address this issue, however the percentage of running injury occasions are still alarming. Studies have yet to develop a good predictive classification model for running injury. Therefore, the goal of this study was to identify the determinants of running injuries, to classify running data by degree of severity and to develop a predictive classification model of running injury. Two case studies related to running injury were retrieved from the public available domain. Data mining approach was conducted to pre-process and to classify data into three injury levels: low, moderate, and severe risks aided by Waikato Environment for Knowledge Analysis (WEKA) version 3.8.6 tool. The J48, SMO, Random Forest, and Simple Logistic algorithms were used for 10-fold cross validation mode classification benchmarked on the ZeroR baseline algorithm. Findings reveal that classification accuracy obtained were from 70% to 100%.

TABLE OF CONTENT

CHAPTER 1: INTRODUCTION

1.1 Overview

This chapter introduces research study related to running injuries and its severity, The chapter is structured into five sections to present the introduction of my research study. The purpose of the study is basically to provide an introduction on the risk factor of running injuries, the injury severity as well as its characteristics, and the high-risk groups of running injuries were highlighted. The project background of this research will be presented. Next, the list of objectives that are to be accomplished are highlighted. Then, the work scope and the project outline of work for this research study are outlined.

1.2 Project Background

One of the most popular recreational sports is running. In the 1970s, running as a form of recreation first spiked in popularity. Even while running is a good kind of exercise, overdoing it might result in lower extremity injury. The overall annual incidence rate for running injuries ranged between 37 and 56% among ordinary leisure runners who were regularly trained and occasionally engaged in long distance runs. [48]

The athletes of the 1970s and 1980s were unlike those of the present day. The subjects in studies conducted in the 1970s and 1980s were serious runners who were thin and predominantly trained by running. They were also mostly (75 percent) male [22] and in their mid-thirties [40,2]. More recent studies have shown that the majority of runners are recreational, completing marathons largely for the sake of finishing, frequently overweight, engaging in cross-training in addition to running, and being in their mid-forties or older [3,4,49]. In the new millennium, there are more female runners (54%) [4]. The populations examined in the various epidemiological studies vary as well; some contain both novice and competitive runners [4,49,44]. Statistics on injury frequency are complicated by these variations in the study populations, making it challenging to understand comparisons between studies.

Statistics have shown that knee injuries typically account for 25% of all running injuries. Other commonly injured body regions are the feet (2-22% of injuries), the ankles (9-20%), lower leg (2-30%), shin (6-31%), upper leg (3-18%), back (3-11 %) and the hip/pelvis/groin (2-11%) [28]. In general, the majority of running injure the lower extremity of the body i.e., knee and bewlo (70-80% of cases). In terms of the nature of injury, 50-70% of running injuries are predominantly musculoskeletal in nature and associated with overuse injury [28].

Overuse injuries are the most common type of injury experienced by runners [1]. As a result of the body's incapacity to heal itself as a result of repeated trauma, they are brought on by repetitive stress to the tissues involved. Overuse injuries in runners are frequently characterised by pain and stiffness. Depending on the severity of the injury, the runner may feel pain and stiffness before, during, or after the run. The runner will eventually stop running if they experience constant pain and stiffness. The causes of overuse injuries in competitive athletes have been examined in several research [16]. The several other musculoskeletal-based injuries are associated with overuse [1]. Running injuries have a clear aetiology that includes aspects relating to the runner and their environment as well as the repetitive nature of the movements required to run. The constant repetition of the exact movements needed to run, as well as factors specific to the runner and their environment, significantly influence the aetiology of running injuries.

1.3 Problem Statement

Overuse injuries are the most frequent sort of reported injuries among runners. Through the use of physiotherapists and nutritionists, numerous safeguards were taken, although overuse injuries were still a possibility. Different models have been created in the past to address the problem of running injuries, but it is unclear what factors are most crucial in determining injury levels. We have not yet created a predictive data mining algorithm that can differentiate between different injury severity levels.

The significance of epidemiology and biomechanics regarding running injuries has attracted a lot of research interest. Existing studies [28] considered the underlying causes and prevention of running injuries. However, there are significantly few thorough, controlled studies testing the efficacy of injury prevention strategies in running.

1.4 Study Objectives

The goals of the study are to

I. Identify the determinants of the overuse and misuse injuries in running motion
II. Classify the running data by degree of severity
III. Develop a predictive classification model of running injury

1.5 Work Scope

This research focused on running parameters for constructing a predictive injury classification model to reduce running injuries. Two cases study involving running injuries among runners were obtained from the public available domain. This research entails the use of data mining analysis tools, which necessitates the use of WEKA (Waikato Environment for Knowledge Analysis) version 3.8.6 software. The data mining analysis is divided into six main stages: Data Collection, Data Preprocessing and Data Extraction, Data Classification, Verification, Model Building and Evaluation. Three classes of running injuries are included: low risk, moderate risk, and high risk. Data is pre-processed to remove any outliers, extreme or missing values. The data categorization method involves data classification analysis at using J48, Random Forest, Simple Logistics, SMO and BayesNet algorithms.

1.6 Project Outline

This thesis is divided into five chapters.

The first chapter presents an overview of running injuries, including their origins and severity levels. The study background, objectives, problem definition, and scope of work are described.

The second chapter contains the literature review of the research on running injuries as well as the factors contributing to the severity of the injury. The important findings from past studies are also presented.

The data mining approaches applied in this study is outlined in Chapter 3. This chapter describes the procedures employed, beginning with data collection followed by data pre-processing and classification analysis. The development of predictive injury classification model is also presented.

The findings of the study are discussed in Chapter 4. The attributes utilised to develop the classification models and the prediction accuracies are displayed in tables and graphs. The categorization results are reviewed, analysed, and compared further. The primary characteristics, particularly the main attributes that distinguish the severity levels of running injuries are underlined.

The fifth chapter summarises and concludes the whole research study. This chapter also discusses the study limitations as well as prospective future works.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview

Recreational and competitive running is practised by many individuals to improve cardiorespiratory function and general well-being. This chapter presents the state-of-the-art review on major negative aspect of running is the high rate of injuries to the lower extremities. Several well-designed population-based research have been studied to find the effect of foot strike pattern and shorter strides on running injuries, running risk injuries assessments and the other demographic factors contributing to running injury. There were altogether 56 related works extracted from Science Direct and Google Scholar search engines reviewed in this chapter.

2.2 Search Strategy

Two search engines, Science Direct and Google Scholar were consulted for eligible articles related to the preventive model for running injuries. The articles were searched by using the following keywords: "preventive model for running injuries", "overuse running injuries" and "lower extremity injuries in running".

The time duration in the search covered was from the beginning of publications until recent year 2022. A total of 4190 articles were identified from the electronic search of the literature, and many article was retrieved from the reference lists in the articles identified which was around 3514 hits (Figure 2.1). A total of 105 duplicates were removed, and 1604 articles were found irrelevant and excluded based on review by titles and abstracts. Of the remaining 1805 hits, 77 articles were excluded due to irrelevant contents based on full-text screening as well as 1130 articles as the full-texts are unavailable. Further refining examination from the 598 articles found 405 articles were excluded based on inaccuracy of articles and topics as it falls outside the scope of study which is about lower extremity injury.

The inclusion criteria are as follows

1. Risk factors & Running Related Injury (RRI)
2. Factors contributing to overuse running injury severity
3. Underlying medical conditions
4. Details of lower extremity affected

The remaining 56 articles are used for literature studies.

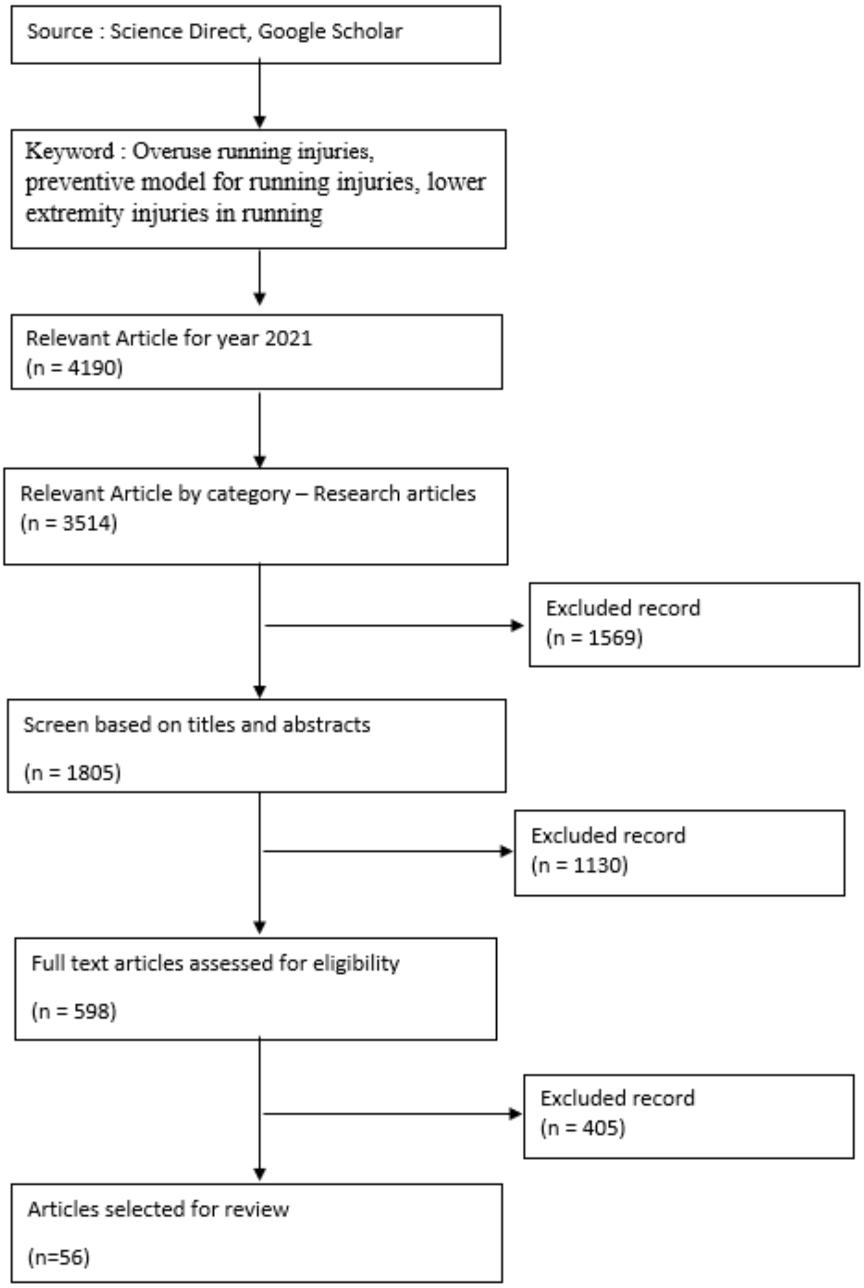

Figure 2.1 : Schematic diagram of search articles

Table 2.1: Summarizes the objectives of 56 related papers and the parameters used in their literature. Detailed analyses are presented in the following sections.

Reference	Purpose of study	Parameters
[1]	To better understand the role of gluteal muscle activity in the aetiology, presentation, and management of PFPS	Gluteus Medius (GMed) activity, patellofemoral pain syndrome (PFPS)
[2]	To study the occurrence of running-related injuries and the distance covered during the training at the start of the training program	Injuries, distance covered
[3]	To study the incidence and the sex-specific predictors of running-related injury (RRI) among a group of recreational runners training	Running-related injury
[4]	To investigate predictors for running-related injuries (RRIS) will differ between male and female novice runners	Running-related injury
[5]	To determine whether the stance phase kinematics and ground reaction forces in running are different between younger- and older-aged men.	Running speed, knee flexion, impact peak force, initial loading rate
[6]	To evaluate the effects of motion control and cushion trainer shoes on running mechanics in low-and high-arched runners.	Three-dimensional kinematic and kinetics, motion control and cushion trainer shoes
[7]	To describe the baseline characteristics of participants starting a 25-week marathon training program, and their relationship to injury risk factors.	Weight, Running-risk injuries
[8]	The use of minimal footwear in other populations besides runners is discussed. Finally, areas for future research into minimal footwear are suggested	Running shoe
[9]	To investigate components of core stability as potential risk factors for acute lower extremity injuries	Acute injury group
[10]	To discover the prevalence of injury to the lower extremity among runners and any associated risk factors which correlate with the development of lower extremity injuries	Injury to lower extremity; knee, calf/lower leg, hip/pelvis/upper leg

[11]	To examine the relationship between navicular, drop and a) rearfoot eversion excursion, b) tibial internal rotation excursion, c) peak ankle inversion moment, and d) peak knee adduction moment during the stance phase of running	Navicular drop measures
[12]	To examine the effects of shoe type and stride length reduction on lower extremity running mechanics and cumulative loading	Measures of loading at the metatarsophalangeal joint and ankle joint, stride length
[13]	To determine if differences in lower extremity and trunk biomechanics during running exist among runners with current ITBS, previous ITBS, and controls	Iliotibial band syndrome (ITBS), Hip abductor strength
[14]	To compare the three-dimensional lower extremity running kinematics of young adult runners and elderly runners.	Age, stride length, movements of the lower extremity
[15]	To determine the incidence of RRI in the lower limbs and spine in a sample of recreational runners, and to determine which training or personal characteristics may be considered predictive factors for RRI in this population	Running-related injury at muscle and knee
[16]	To investigate which preventive measures runners use when preparing for a half- or full-marathon and whether the use of these measures at baseline and during the preparation-period differs between runners who sustained no/non-substantial running-related injuries (NSIRS) or substantial running-related injuries (sirs	Running shoe, Running-related injury, substantial running-related injuries
[17]	To record the prevalence of health problems in recreational long-distance runners preparing for an event.	Weekly mileage, Age, Health problems
[18]	To conduct a systematic literature search and synthesize the evidence about the relationship between foot posture and running biomechanics	Foot posture and subtalar joint kinematics
[19]	To review the current state of knowledge related to overuse running injuries, with a particular emphasis on the effect of impact forces	Overuse injuries of lower extremity
[20]	To identify biomechanical and anthropometric variables that contribute to overuse injuries in runners	Hamstring flexibility
[21]	To (i) investigate whether chronic endurance running is a sufficient stimulus to counteract the age-related changes, (ii) identify adaptational phenomena in running mechanics due to	Muscle–tendon units (MTUS), age, running mechanics

	age-related changes, (iii) examine whether chronic endurance-running exercise is associated with adaptational effects on running characteristics in old and young adults	
[22]	To suggest causes and solutions based on the 'fundamental principles' approach that characterises engineering	Risk factors for running injury
[23]	To study how runners coped with the impact caused by the foot colliding with the ground barefooted and with modern shoes	Kinematic and kinetic analyses, impulse, the body's momentum
[24]	To determine if measurements of static lower limb alignment are related to lower limb injury in recreational runners	Lower limb injury, lower limb alignment
[25]	To study lower extremity injuries in runners	Lower extremity injuries
[26]	To investigate how to predict lower extremity injuries among habitual runners	Lower extremity injuries
[27]	To investigate the association between training-related characteristics and running-related injury using a new conceptual model for running-related injury generation	Running-related injury
[28]	To study running the race against the injuries	Running related injury
[29]	To determine the risk factors that differentiate recreational runners who remain uninjured from those diagnosed with an overuse running injury	Mental health–related quality of life, knee stiffness, sex
[30]	To systematically review evidence for kinematic risk factors for (lower limb tendinopathy) LLT in runners.	Knee internal rotation, rearfoot eversion, knee flexion at heel strike
[31]	To determine the association between hip abduction strength and lower extremity running related injury in distance runners.	Iliotibial band syndrome, patellofemoral pain syndrome, medial tibial stress syndrome, tibial stress fracture, Achilles tendinopathy
[32]	To investigate foot posture (measured statically) as a potential risk factor for lower limb overuse injuries	Pronated foot posture
[33]	To assess differences in hip strength, iliotibial band length, and hip and knee mechanics during running between male runners with iliotibial band syndrome (ITBS) and healthy controls	Hip internal rotation, knee adduction

[34]	To study cumulative loads, increase in the knee joint at slow-speed running compared with faster running	Cumulative load, speed
[35]	To determine foot kinematic patterns during running among novice runners with pronated feet and the presence of a relationship between these foot kinematic patterns and foot muscle morphology	Navicular height (NH) at initial contact, dynamic navicular drops (DND)
[36]	To examine the current literature on running related injuries and stipulate whether these can be mitigated through proper footwear selection.	Running related injury, Footwear
[37]	To assess the relationship between running-related injuries, foot posture and other factors in novice runners.	High supination, running surface, number of shoes used, body mass index
[38]	To determine whether hip abduction strength asymmetry exists in female runners with early unilateral patellofemoral pain syndrome (PFPS)	Hip abduction strength
[39]	To study factors associated with a history of plantar fasciitis in female runners	Plantar fasciitis
[40]	To investigate the effects of frequency and duration of training on attrition and incidence of injury	Oxygen volume, treadmill performance, heart rate
[41]	To study leg-length inequality and running-related injury among high school runners	Stride length
[42]	To observe differences in the kinematic profiles of healthy runners (CON) and runners with mid-portion Achilles tendinopathy (ATG)	Sub-talar joint eversion displacement, dorsiflexion velocity, frontal plane ankle joint range of motion
[43]	To examine knee and ankle loading in barefoot and barefoot inspired footwear in relation to conventional running shoes.	Patellofemoral kinetic
[44]	To determine the number of injuries that occur in a running programme designed to minimise the injury rate for athletes training	Age, body mass index
[45]	To investigate the relationship between functional and static foot posture and medial tibial stress syndrome in distance runners	Medial tibial stress syndrome

[46]	To review information about risk factors and sex-specific differences for running-induced injuries in adults.	Sex
[47]	To present a systematic overview of published reports on the incidence and associated potential risk factors of lower extremity running injuries in long distance runners	Lower extremity running injuries
[48]	To study the importance of complete rehabilitation and the early recognition of symptoms of overuse, and on the provision of training guidelines.	Lower extremity injuries, Aetiological factors
[49]	To describe the prevalence and incidence of lower extremity injuries occurring before and during the Rotterdam marathon, and to evaluate the impact of the injuries	Incidence of lower extremity injuries
[50]	To identify risk factors for lower extremity injuries in male marathon runners	Risk factors for lower extremity injuries
[51]	To review information about risk factors for lower extremity running injuries in both short-distance and long-distance runners	Types of runners, Higher body mass index, higher age, sex (male), having no previous running experience, and lower running volume
[52]	To describe the self-reported injury, training, and running technique choices of regular runners in four international regions	Training and footwear choices
[53]	To identify the incidence of Medial tibial stress syndrome (MTSS) and tibial stress fracture (SF) in high school runners and to determine risk factors	BMI, internal rotation of the hip
[54]	To examine many forms of stretching available and evaluates what athletes can and should do	Types of stretching
[55]	To discuss the repetitive trauma that overwhelms the tissue's ability to repair itself	Isometric strength of the hip flexor, extensor, abductor, adductor, and external and internal rotator muscle
[56]	To investigate whether chronic endurance running is a sufficient stimulus to counteract the age-related changes in the mechanical and morphological properties	Maximal eccentric hip abduction strength, Distance

2.3 What is Running and Overuse Injuries?

Running which is commonly known as a kind of exercise was defined in various ways. According to McGrath et al. [28], Running is a gait pattern that resembles a spring because it comprises two flight phases with synchronised potential and kinetic energy and the centre of mass over a compressed support leg at mid-stance. Humans naturally transition from walking to running as their speed rises [28]. Recent studies have also revealed that lower limb discomfort brought on by higher stresses as velocity increases is the catalyst for changing from walking to running.

The most often reported ailment among runners is overuse injury. Such injuries are caused by the tissues being subjected to repeated stress, which overwhelms the body's natural ability to repair itself [28].

According to Baxter et al. (2016) [54], overuse injuries in runners typically start out with discomfort and stiffness. Depending on the severity, the runner may experience pain and stiffness before, during, after, or a mix of these during the run. Running will ultimately come to an end due to ongoing discomfort and stiffness. Competitive athletes have been the focus of the majority of studies on overuse injury risk factors [54]. Repeated stress to the affected tissues—bone, ligaments, or tendons—leads to an overuse injury as well [54]. Overuse injuries can affect a variety of tissues and anatomical locations, but the underlying cause is always the same: repeated damage overwhelms the body's capacity for self-healing [54]. The forces and repeated nature of the gait cycle may be related to this. Alternatively, a past injury that the body tried to make up for by putting more stress on another area of the body could have caused an overuse injury [28]. This would eventually cause tissue breakdown and obvious injury in the vulnerable region.

2.4 Factors contributing to running injuries

2.4.1 Foot strike pattern and shorter strides

Foot strike is another term for foot placement, which refers to how your foot impacts the ground when running according to (Guide To Proper Running Form Foot Strike, n.d.). Every step includes a combination of foot landing and rolling. Every runner has an own running style (and therefore foot placement) [59]. This reflects your running technique, as well as your footwear and terrain. There are three sorts of foot strikes which are rear foot strikes (RFS), fore-foot strike (FFS) and mid-foot strikes (MFS). What section of your foot falls first in your stride determines whether you do one or the other. Rear foot strike refers to landing on your heel first in your stride. Mid-foot strike means landing in your stride with the middle of your foot first. The foot initially strikes the ground in the posterior third of the length of the foot with a rearfoot strike. Strides is to go in a certain direction with lengthy, deliberate steps.

Lieberman et al. [23] examined the running related injuries and stipulated whether these can be mitigated through proper foot strike pattern selection. The authors have considered both engineering and medical aspects to provide insights into the underlying impact mechanics and injury precursors. Such research might help to improve our understanding of injury risk and prevention. [19, 36]. Accordingly, running can be most injurious at the moment the foot collides with the ground [23]. This collision can occur in three ways: a rear-foot strike (RFS), in which the heel lands first; a mid-foot strike (MFS), in which the heel and ball of the foot land simultaneously; and a fore-foot strike (FFS) as explained in previous paragraph. Besides, the use of orthotic inserts in shoes and hip abductor weakness were found to be associated with an increased injury risk [31,46].

Firminger et al. [12] attempted to analyze how minimalist shoes and shorter strides affected lower-extremity mechanics and cumulative loads. Running in minimalist shoes was linked to higher single-stance loading at the knee joint but decreased single-stance loading at the metatarsophalangeal (MTP) and ankle joints. The (MTP), ankle, and knee joints are evaluated for peak moments, angular impulse, mechanical work, and cumulative impulse in the sagittal plane [43]. On the other hand, running with a shorter stride length was associated with reduced single-stance assessments of ankle and knee joint loading. Cumulative stresses increased in the MTP and ankle joints during shorter strides but decreased in the knee joint [20]. Even though both load reduction strategies seemed to have an additive effect, reducing knee joint stresses with a shorter stride length was more effective than doing so with a minimalist shoe. [34]. According to (Bonacci et al., 2013), running in a minimalist shoe raised joint stresses at the MTP and ankle joint, which could explain why minimalist shoe users have a higher rate of overuse injuries. Ankle and knee loads were lowered by 10% by shortening the stride, however only the cumulative loads on the knee were decreased [12].

It is vital to identify foot kinematic patterns that lead to injury during running to prevent RRIs. Okamura et al. [35] defined foot kinematic patterns during running in new and young runners with pronated feet who are at high risk of RRIs, as well as to characterise foot muscle morphology in each pattern. RRIs are induced by overuse as a result of frequent running [30]. Many runners who suffer from running-related injuries experience plantar fasciitis, medial tibial stress syndrome, and iliotibial band syndrome Kakouris et al., 2021. As a result, RRIs have been linked to aberrant lower limb kinematics during running, which puts more load on the tissues. Foot kinematics during running is a significant component. Increased foot pronation during running, which indicates a lower medial longitudinal arch (MLA), is a major risk factor for a variety of RRIs, including patellar tendinopathy and posterior tibial tendon dysfunction [39].

Furthermore, determining the morphology of foot muscles in each pattern aids in the ddevelopment of appropriate preventive activities. The presence of a supinated or pronated foot was related with high and substantial risks when compared to feet in a neutral standing position as per Pérez-Morcillo et al. [37]. Tibial stress syndrome and Achilles tendinopathy were found to be substantially linked to pronation [42, 45].

There is also considerable evidence of a link between non-neutral foot types and lower extremity injuries, according to systematic reviews, however the amount of the effect appears to be small Tong & Kong, 2013. The most prevalent foot type is neutral. During the gait cycle, a runner with a neutral foot type falls on the heel and rolls forward until the impact is uniformly dispersed across the forefoot. When compared to a neutral foot, a severely supinated foot has the highest risk of RRI in inexperienced runners [32]. When compared to a neutral foot, moderate supination and pronation were both linked to an elevated risk of RRI.

Buldt et al. [6] studied the association between foot position and lower limb biomechanics in walking using a systematic method. Karsten Hollander et al. [18] attempted to review how different foot positions influenced running biomechanics. The features of the foot and medial longitudinal arch have repeatedly been brought up as possible risk factors for lower extremity overuse injuries [18]. High arched feet are supposed to be stiffer with less ability to absorb stress, while the research is not convincing. Ankle injuries, bone injuries (particularly to the tibia or femur), and injuries to the lateral side of the lower leg may become more common as a result. In contrast, low arched feet seem to be associated with increased knee injuries, soft tissue injuries, and lower limb injuries on the medial side [18]. Due to anatomical restrictions, a relationship between the medial longitudinal arch and subtalar joint kinematics and tibial rotation can be assumed [11]. In terms of biomechanics and injury epidemiology, the interaction of foot type and footwear is likely only one component of the multifaceted aetiology of running-related injuries. The relationships found in this comprehensive review, with the exception of rearfoot eversion excursion, are too varied to provide evidence or clinical guidance [17].

2.4.2 Age and Gender Effects

Fukuchi et al. [14] evaluated the three-dimensional kinematics of the lower extremity during the stance phase of running in young adult runners versus senior runners [14]. Their findings showed that the senior have a higher rate of injury and a longer time to return to running. This could be attributable to deterioration of the musculoskeletal system as age increases, as well as changes in running movement patterns between senior and young adult runners. Besides, muscle weakness, changes in the mechanical characteristics of collagenous tissues, and changes in muscle architecture are all linked to the ageing process.

The age-related changes in the musculoskeletal system, such as increased joint stiffness and reduced function of the triceps surae and quadriceps femoris muscle-tendon units have also been observed in Karamanidis & Arampatzis [21]. Elderly runners had a higher knee flexion at heel strike and a smaller knee flexion/extension range of motion than the young adults [5]. The authors found no linkage between decreased knee flexion/extension range of motion and stride length within or between groups. The senior runners have a reduced knee joint range of motion and shorter stride length owing to an increase in ankle and knee joint stiffnesses as they get older, or to the muscle-tendon units' weakening [21]. This explains why elderly individuals could be more susceptible to running-related injuries.

During the stance phase of running, senior runners have a lower extremity movement pattern that differs from that of young adult runners [14]. Such aspect explains why elderly people are more prone to running-related injuries. Additionally, the increased impact loads in older runners under controlled running speed settings point to a reduction in the musculoskeletal system's capacity to absorb shock, which in turn increases the stress on bone, joints, and soft tissue in the lower extremities.

Jade Dempster et al (2021), hypothesized that there is a significant incidence of lower extremity injuries both sexes of runners, with female runners suffering more injuries than male runners. [10]. Female runners were found to have a greater injury rate (39.7%) than male runners (34.3%) [25]. The knee was the most frequently injured anatomical location overall, however the prevalence of running related injuries at anatomical locations varied between genders [10]. Males were more likely to be injured in the hip/pelvis/upper leg area, whereas females were more likely to be injured in the lower leg/calf area [10]. In contrast, for females, the least common area for incidence of RRI was within the ankle/foot area. Further, the results from [25] also indicated that the male runners least injured area was the calf/lower leg, despite the female runners presenting the greatest proportion of RRI in this region.

High school male and female cross-country runners, according to Rauh, M. J. (2018), show that leg-length disparity was not connected with RRI, with the exception that boys who had a leg-length inequality > 1.5 cm were more likely to experience lower leg (shin/calf) RRI. Further, the shorter or longer limb was not associated with side of RRI [41]. Past studies have shown comparatively lack of association between the injured side and the side of the leg-length inequality in this research.

2.4.3 Muscular aspects

The hip abductors, a crucial set of muscles, keep you steady and powerful throughout each step during jogging. These muscles on the lateral (or outer) side of the hip assist your stride and aid in your pain-free mobility. Thijs et al. [55] examined the relationship between hip abductor weakness and Patellofemoral pain syndrome (PFPS), however they could not discover a connection. Ramskov et alstudy [56].'s computed risk development scores based on cumulative run-distance. According to previous research, people with weak hip abductors are more likely to develop PFPS at 25 and 50 cumulative kilometres of training [56] than people with stronger hip abductors. After running a total of about 100 km, this relationship started to decline. As a result, it is challenging to tell whether hip abduction weakness is predictive or associative with PFPS. Weakness of the hip abductors and external rotators is suggested to result in excessive femoral adduction and internal rotation, thereby increasing the valgus vector at the knee during dynamic tasks such as running.

Caldwell et al. [31] compiled all the existing evidence related to hip abduction strength and described the link between hip abduction muscle strength and running-related lower extremity injury in distance runners. Hip abduction strength has been shown to improve knee joint kinematics by lowering knee valgus and improving patellar tracking in male runners, whereas hip abduction weakness has been linked to altered transverse and frontal plane kinematics [13,33]. When investigated in conjunction with iliotibial band syndrome (ITBS), it revealed that hip abductor musculature strength had a stronger link to injury. Yagi et al. (2012) [53] only discovered a connection between hip abduction strength and TSFx in female high school runners. The BMI did not differ significantly across the groups, but there were variations in hip rotation angles, which were associated with injury.

Cedric De Blaiser et al. (2021) investigated core stability components to identify modifiable core stability-related risk factors for the development of acute lower extremity injuries [9]. Modifiable risk factors for the development of non-contact acute lower limb sports injuries include decreased core musculature strength and endurance, increased strength asymmetry, decreased core proprioception and neuromuscular control, and poor postural control. Handheld dynamometry revealed that increased hip strength asymmetry for isometric abduction strength was an organically changeable risk factor [38]. Other probable compensating mechanisms for hip abduction strength imbalance during unilateral training have also been described. A frequent compensatory strategy during locomotion is to elevate the contralateral pelvis and tilt the trunk towards the comparatively weak abductors. The ensuing ground response force vector is closer to the hip joint centre, reducing the burden on the weaker hip abductors.

2.4.4 Experience

Mucha et al. [31] concluded that the aetiology of running injuries is clearly influenced by the continual repetition of the identical movements required to run, as well as elements relating to the runner and their environment. Short- and long-distance runners have several risk factors for lower extremity injuries, however evidence for these factors in running-related injuries is weak. Both short- and long-distance runners appear to suffer from multifactorial running ailments [31]. Several theories have been postulated as to the etiology of running related injury (RRI) including impact forces [19], foot pronation [8], foot strike pattern [23] and hip muscle function [1].

The training strategies chosen by the runners could be a cause for the variations in injury risk variables. Runners are also confronted with a multiple assortment of training options that have been linked to injury risk, running experience, volume, intensity, and regularity. [47, 52]. Running more than three times a week, having sustained an injury in the preceding twelve months, and having less than five years of running experience were all common risk factors for developing running-related ailments [29,41].

Overuse running injuries are less likely to occur in runners who use a running stride with relatively low impact forces and a moderately rapid rate of pronation [20]. Because impact forces grow with speed, it would be sensible to advise injured runners to reduce their training speed to lessen impact forces.

The incidence of RRI in recreational runners must be determined to track interventions that may have an impact on the rate of RRI in this population [3]. Previous RRI was linked to a higher risk of RRI among recreational runners [3]. Previous injuries, running more than 64 km/week, and fewer than three years of running experience have all been recognised as risk factors for RRI in this cohort.

Higher RRI was also linked to longer training sessions, faster training, and interval training. Speed training was associated with higher RRI [15]. This can be explained by an increase in the running intensity overloading the musculoskeletal structures, predisposing recreational runners to injury [15]. The majority of recreational runners who engage in interval training switch from normal or slightly higher intensity intervals to lower or much lower intensity intervals (e.g., walking), resulting in a lower total training intensity in a given running session and a lower risk of injury.

Body mass index and previous injury determine the impact of a runner's training load on running-related injury [27]. His findings highlight the need of distinguishing between confounding and effect-measure modification in injury research involving runners. Malisoux et al. [27] proposed a conceptual model of RRI generation in which training-related properties are the major exposures of interest to running injuries [27]. Because the effect of exposure to training-related features varies across strata of non-training-related parameters, non-training-related qualities are thought to be plausible effect-measure modifiers (EMM) [1]. A positive synergy exists between prior injury and weekly volume and session frequency, whereas a negative synergy exists between BMI and weekly volume.

To avoid RRIs, van Poppel et al. (2021) said that additional research is needed not just on probable injury risk factors but also on effective preventative interventions. To far, studies have concentrated on various measures, including as stretching, warming-up, and cooling-down, as the primary emphasis RRIs, but none have been found to impact the occurrence of RRIs specifically in marathon runners [7, 16].

Alan Hreljac et al. (2000) evaluated runners' risk of lower extremity overuse injuries by identifying biomechanical stride characteristics and anatomical features that may predispose a runner to overuse injuries, as well as attempting to control for training variables by matching groups based on several important training factors.

2.5 Summary

There were not many local data as Malaysians have not ventured into the field of lower extremity injury due to running. According to the results of the research, prior injury is still the most important risk factor for RRIs. Despite the fact that there is conflicting data in the literature, it is still necessary to validate that several training features appear to be at play. Clinicians could better understand RRIs and how to avoid them with the aid of more study. There are a number of lower extremity injury risk variables that have been linked to runners, although the quality of the supporting data is sparse.

CHAPTER 3: METHODOLOGY

3.1 Overview

This chapter focuses on data mining approach to study the relationship between the running motion and the severity of the running injuries to gain better understanding of this research topic. The entire study involves data collection, data pre-processing, data classification and model building as well as verification of the model. Data pre-processing effort detects and removes potential data outliers and extreme values. At this level, data imputation method was used to treat missing values. In order to classify data by injury severity levels: mild, moderate, and severe, the J48, SMO, Random Forest, and Simple Logistic algorithms of WEKA are used. The performance of the algorithms is examined on percentage classification accuracies. A classification predictive model that relates to running injuries factors such as distance covered during running, time taken, intensity and the other attributes is built to classify running injuries severity.

3.2 Approaches

This study applies eight stages of data mining approach to develop a predictive classification model for running injury. The flow chart of the entire data mining process is shown in the figure below.

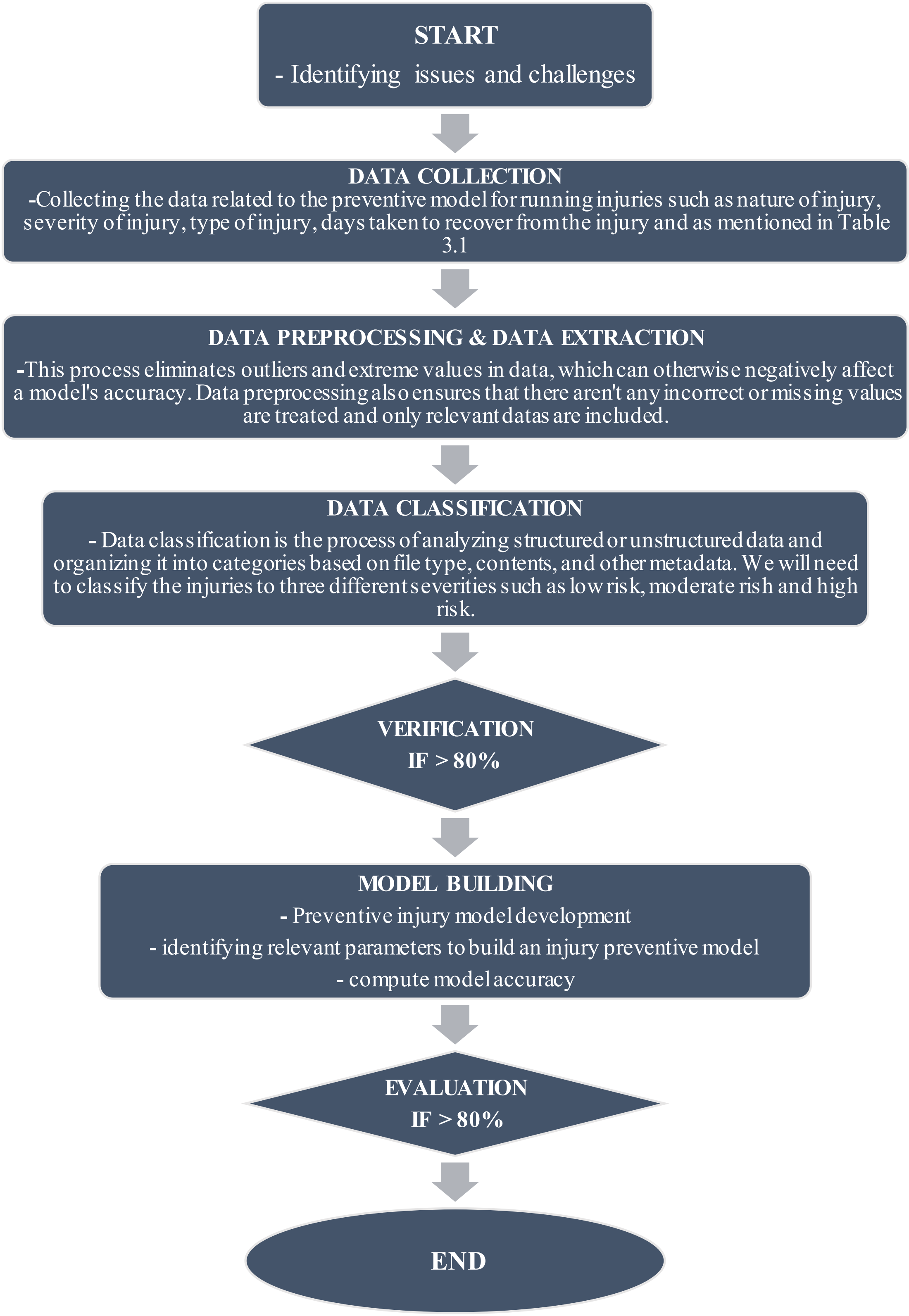

Figure 3.1 : Flow Chart of Data Mining Process