USING NEURAL NETWORKS TO ESTIMATE LOWER BODY KINEMATICS FROM IMU DATA IN JAVELIN THROWING

Krista Vohlakari

Master’s thesis of biomechanics
Faculty of Health and Sport Sciences
University of Jyväskylä
Spring 2022
Supervisors: Janne Avela, Neil Cronin,
Timo Rantalainen & Miika Köykkä

TIIVISTELMÄ

Vohlakari, K. 2022. Neuroverkkojen käyttäminen alavartalon kinematiikan ennustamiseen inertiaalisensoridatasta keihäänheitossa, Liikuntatieteellinen tiedekunta, Jyväskylän yliopisto, biomekaniikan pro gradu -tutkielma, 61 s.

Tämän opinnäytetyön tarkoituksena oli hyödyntää koneoppimista kinemaattisten muuttujien arviointiin inertiaalisensoreista saatavasta datasta keihäänheitossa. Perinteisten liikeanalyysimenetelmien rinnalle olisi hyvä löytää menetelmiä, jotka mahdollistaisivat nopean palautteen antamisen urheilijoille. Datankeruun helpottuessa myös aiempaa tarkempi ja suuremmalle kohderyhmälle tehtävä tutkimus mahdollistuisi.

Kymmenen tavoitteellisesti korkealla tasolla kilpailevaa keihäänheittäjää osallistui tutkimukseen. Mittaukset järjestettiin Kuortaneen urheiluopistolla sisätiloissa keväällä 2021. Urheilijat suorittivat tutkimusprotokollan osana tavallista harjoitteluaan. Suoritukset mitattiin optoelektronisella Vicon-järjestelmällä. Tämä menetelmä toimi ”kultaisena standardina” muiden menetelmien vertailua varten. Vicon-järjestelmään yhdistettiin Bluetoothin avulla langattomasti BlueTrident-inertiaalisensorit, jotka kiinnitettiin heittäjän lantiolle, reisiin ja sääriin.

Viidestä sensorista saatuja lineaarikiihtyvyyksiä ja kulmanopeuksia käytettiin pitkäkestoista työmuistia hyödyntävien LSTM-neuroverkkomallien syötteinä. Vastemuuttujia olivat lantion orientaatio (tilt, list, rotation) sekä lonkkien ja polvien kolmiulotteiset nivelkulmat. Ohjattua oppimista varten vastemuuttujat laskettiin Vicon-järjestelmästä saaduista markkerien koordinaateista avoimen lähdekoodin OpenSim-ohjelmistolla. Kolme erilaista mallia koulutettiin. Yksi hyödynsi jokaista viittä sensoria ja opetus tehtiin jättäen vuorotellen yksi urheilijoista testidataksi. Toiseen malliin käytettiin ainoastaan tukijalan reisi- ja säärisensoria ja vastemuuttujana oli pelkkä polvikulma. Kolmas malli käytti kaikkia viittä sensoria, mutta se oli urheilijakohtainen ja opetus tehtiin jättäen vuorotellen yksi maksimaalinen heitto testidataksi.

Urheilijakohtainen malli ennusti kaikki muuttujat parhaiten (RMSE, 2–5.4 astetta; ICC1, 0.88–0.98), mutta oli luultavasti ylisovittunut datan vähyyden vuoksi. Viiden sensorin ryhmämallit ennustivat muuttujat RMSE:n ollen keskimäärin 7.7–16.6 astetta eri muuttujissa. Ennustetut taaemman jalan polvikulmat ja lantion kallistuminen sivuttaissuunnassa vastasivat huonosti ”kultaisen standardin” antamia lukemia sisäkorrelaatiokertoimella arvioituna (ICC1). Muut muuttujat pystyttiin ennustamaan keskimääräisesti tai hyvin (ICC1, 0.72–0.87). Kahden sensorin polvikulmamalli oli parempi kuin viiden sensorin ryhmämalli ja huonompi kuin urheilijakohtainen malli. Ryhmämallit toimivat yllättävän hyvin datan vähyyteen verraten. Mikäli dataa ohjattua oppimista varten olisi tarpeeksi, voisi olla kannattavaa testata tiettyihin keihäänheittosuorituksen vaiheisiin perustuvan LSTM-mallin yleistyvyyskykyä.

Asiasanat: kinematiikka, keihäänheitto, tekoäly, koneoppiminen, inertiaalisensori

ABSTRACT

Vohlakari, K. 2022. Using neural networks to estimate lower body kinematics from IMU data in javelin throwing, University of Jyväskylä, Master's thesis, 61 pp.

The aim of this thesis was to use machine learning to predict kinematic variables in javelin throwing using IMU data. In addition to traditional motion analysis methods, new more flexible methods would enable faster feedback to athletes. If data collection could be simplified, it would be possible to collect accurate data from larger populations.

Ten well trained international and national level athletes participated in the study. Measurements were conducted at Kuortane Olympic Training Center in spring 2021. Athletes performed the study protocol as part of their general training. Javelin throw performances were captured via an optoelectronic Vicon system, which was used as the "gold standard" for comparisons to other methods. BlueTrident inertial measurement units were connected wirelessly via Bluetooth to the Vicon system. Sensors were mounted to the pelvic area near the sacrum, thighs and shanks.

Linear accelerations and angular velocities from five sensors were used as inputs for long short-term memory (LSTM) models. Pelvic orientations (tilt, list, rotation) and bilateral hip and knee angles were used as target variables. For supervised learning, the target variables were calculated with open source OpenSim software using marker trajectories recorded via Vicon. Three different LSTM models were trained. The first used data from all five sensors and training was implemented using the leave one out method. The second model used only two sensors mounted to the brace leg, with the goal of predicting only the knee angle of this leg. The third model used all five sensors but was athlete-specific, and training was again implemented using the leave one out method.

Athlete-specific individual models yielded the most accurate predictions for all variables (RMSE 2–5.4 degrees, ICC1 0.88–0.98) but these models were probably overfitted due to the small training dataset. For the five sensor group model, the RMSEs were 7.7–16.6 degrees for all variables. Based on ICC1, the knee angles of the rear leg and pelvic list corresponded poorly with the "gold standard". Other variables were predicted with moderate or good accuracy (ICC1 0.72–0.87). The two sensor knee angle model performed better than the five sensor group model but worse than athlete-specific individual models. Group models worked surprisingly well considering the small dataset. If a larger dataset were available for supervised learning, it would be valuable to explore the generalisability of the LSTM model for predicting biomechanical parameters at certain phases of the javelin throw.

Key words: kinematics, javelin throwing, artificial intelligence, machine learning, inertial measurement unit

LIST OF ABBREVIATIONS

3D	3-dimensional
AHRS	Attitude and heading reference system
ANN	Artificial neural network
CNN	Convolutional neural network
FFNN	Feedforward neural network
IK	Inverse kinematics
IMU	Inertial measurement unit
LSTM	Long short-term memory
ML	Machine learning
NN	Neural network
RMSE	Root mean square error
RNN	Recurrent neural network
SSC	Stretch-shortening cycle

CONTENTS

1 INTRODUCTION

The aim of this work is to explore and test different neural network (NN) models in estimating kinematics in javelin throwing events to get a deeper understanding of implementing machine learning (ML) algorithms for this purpose. Javelin throwing related movement is very individual (Bartlett & Best 1988) despite the fact that the same general rules apply to all throwers. In order to achieve the goal of maximising throwing distance, kinetic energy must be transferred from the thrower to the javelin with as large an impulse as possible (Bartlett & Best 1988). Javelin throwing has been studied for several decades, but traditional motion analysis methods are still generally used for biomechanical analyses of competition performances.

Competitions are good events to measure and study biomechanics of throwing because it is in this setting that throwers seek to achieve their maximal performance. Most javelin throw research has been done using videos captured from competition situations (e.g. Chen et al. 2020; Panoutsakopoulos et al. 2016; Liu et al. 2014; Campos et al. 2004; Whiting et al. 1991). This means mostly 3-dimensional (3D) analysis done using two video cameras. Video based motion analysis is a good method to use in javelin throwing studies because it takes into consideration requirements in competition events. The use of multiple cameras and markers attached to the thrower could improve the accuracy of motion analysis, but traditional motion analysis with body mounted markers cannot be used in competitions, and even if it were allowed, the preparation time would be excessive and impractical. Moreover, the use of multiple-camera opto-electronic system (which could also enable automated processing) is not possible due to environmental constraints.

Analysing one throw manually using video captures from two cameras is time consuming. Thus, it is not possible to give fast feedback to athletes with this method in competitions or training scenarios. Individual feedback is usually based mostly on visual observations by coaches. There is strong demand for fast feedback. Despite the high correlation between release speed and throw distance (Bartlett et al. 1996, Panoutsakopoulos & Kollias 2013; Manesh & Dr. Dhinu 2016; Panoutsakopoulos et al. 2016), elite level athletes need an individualized approach to find out how to achieve the highest possible release speed (Whiting et al. 1991; Mahmud 2009).

Possible solutions to fast feedback systems could be markerless motion analysis, wireless wearable devices or a combination of these. Markerless motion analysis could be done using the same cameras as with traditional motion analysis but with digitization using ML based methods. One weakness of markerless motion analysis is related to environment conditions. Markerless motion analysis is likely more effective with optimal and controlled lighting. Thus, it is easier to implement indoors than outdoors. To get 3D information, the cameras must also be calibrated in the same way as in traditional motion analysis. Ideally, a markerless system could be located indoors, allowing the cameras to be kept in the same place permanently.

An advantage of wearables is that environmental conditions do not have such a big effect with regard to lighting. Wearables are usually accelerometers, gyroscopes or a combination of these two devices, i.e. inertial measurement units (IMUs). These devices measure linear acceleration, angular velocity, and strength of magnetic field, and can be used to estimate segment orientations by implementing appropriate algorithms (Brouwer et al. 2021). However, this method is not straightforward and requires previous domain knowledge. For example, even if the location and body position of the thrower at the beginning of the throw is known, sensors orientation estimates are prone to drift during long measurement times (King 1998). Thus, the determination of the location and the orientation of the sensor is distorted. Complex algorithms must be used to correct the drift error (e.g. Caruso et al. 2021).

ML is also utilised in combination with wearable devices (Mundt et al. 2021). Previous studies have shown that kinetic and kinematic parameters can both be estimated with ML in walking and running studies with promising accuracy (e.g. Sharifi Renani et al. 2021; Rapp et al. 2021). An advantage of ML is that it could enable avoiding the use of complex unstable algorithms which also requires the measurement of the starting position and accurate attachment of sensors to the body segments (Hernandez et al 2021). However, one challenge is lack of data. To be able to develop valid, generalisable ML models, a large volume of data are needed. The appropriate amount of data cannot be determined with formal rules but when the complexity of model increases more data is required (Halilaj et al. 2018).

The aim of this work is to examine which kind of ML methods could be used to estimate kinematics in javelin throwing using data measured with IMUs. This examination focusses on the use of existing approaches, with an emphasis on walking and running. In the research part of this thesis ML models will be built and used to examine which kind of results can be achieved with a small volume of data from ten participants, and how these results could be interpreted.

2 JAVELIN THROWING TECHNIQUE

To examine javelin throwing performance, let us first split the performance into sections: the run-up, the crossovers, the delivery, and the recovery (Bartlett & Best 1988; Morris & Bartlett 1996). According to Ljubisa et al. (2019), researchers have demonstrated multiple ways to distinguish between the different phases. However, the performance starts with the run-up phase to achieve the initial velocity before the most important phases of acceleration (Morris & Bartlett 1996). After the run-up, the athlete performs multiple crossover strides to optimise the position of the javelin. Usually, only the last crossover stride has been studied (figure 1). The landing from the last crossover stride, i.e. the final right foot touchdown (right-handed throwers), marks the start of the delivery phase (figures 1 and 2). (Morris & Bartlett 1996.)

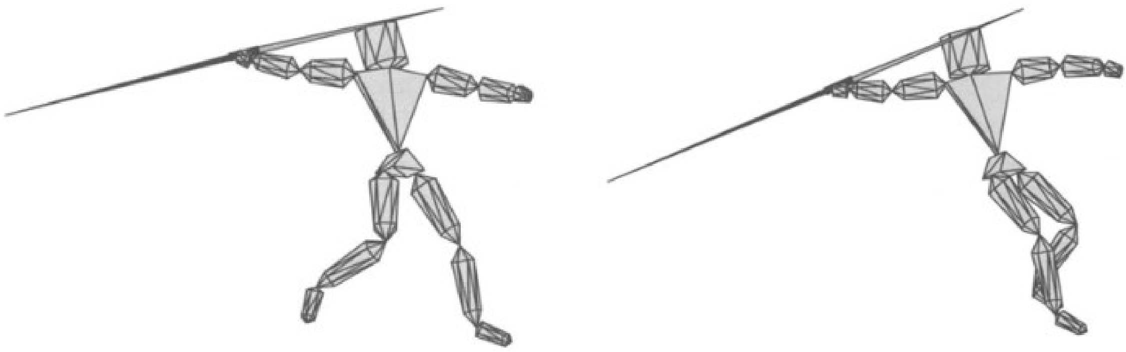

FIGURE 1. The beginning of the last crossover stride and the beginning of the delivery stride (Morris & Bartlett 1996).

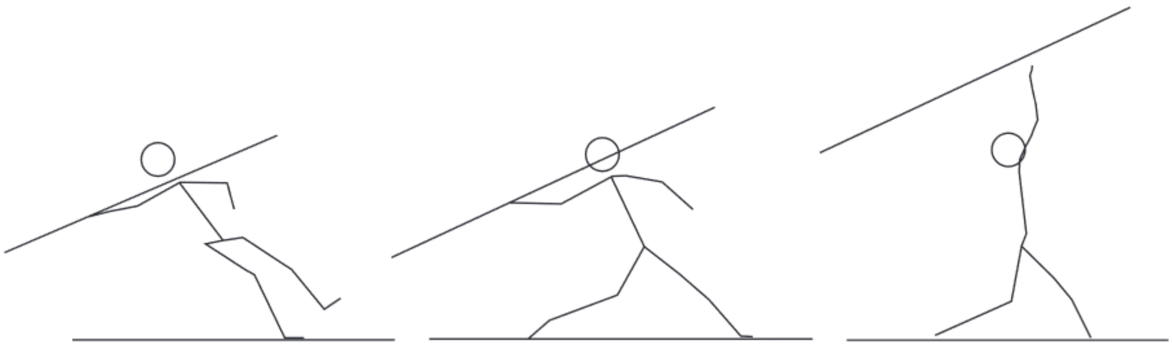

FIGURE 2. The delivery or release phase (Whiting et al. 1991) including the final rear foot contact (beginning of the delivery stride), a brace leg contact, and the javelin release (Campos et al. 2004).

The first sub-phase of delivery can be referred to as the preparatory phase which is the period from the final rear foot (right foot for right-handed throwers) contact to the final front foot contact (Campos et al. 2004). According to the findings of Campos et al. (2004), there is more variation between athletes in the duration of this phase than in other phases. The final front foot contact (later referred to as the brace leg contact) lasts only slightly over 0.1 seconds and has high impact loading (Komi & Mero 1985). It is the time-span when energy is transferred from the body to the javelin before the release of the javelin (figure 2) (Morris & Bartlett 1996). The importance of brace leg contact is not limited to the energy transfer to the javelin. At the end of the throw, after the javelin has been released, the thrower must not step over the foul line which increases the importance of brace leg action (Morris & Bartlett 1996).

Considering the presented phases, the literature has highlighted five phenomena related to javelin throwing performance: the length of the acceleration path of the javelin (Panoutsakopoulos & Kollias 2013; Bartlett et al. 1996; Best et al. 1993; Whiting et al. 1991), utilisation of the stretch-shortening cycle (Campos et al. 2004; Mero et al. 1994; Whiting et al. 1991; Komi & Mero 1985), the sequence of the peak velocities of the joint centres and segments (Liu et al. 2014; Liu et al. 2010; Campos et al. 2004; Bartlett et al. 1996; Mero et al. 1994; Best et al. 1993; Whiting et al. 1991), events related to the brace leg contact (Campos et al. 2004; Bartlett et al. 1996; Mero et al. 1994; Best et al. 1993; Komi & Mero 1985), and individuality (Manesh & Dr. Dhinu 2016; Liu et al. 2014; Liu et al. 2010; Campos et al. 2004; Mero et al. 1994; Best et a. 1993; Whiting et al. 1991). The next subchapters consider examples related the above-mentioned circumstances.

2.1 The acceleration path of the javelin

The acceleration path is the route the javelin travels from the end of last crossover stride until the release of the javelin (Morris & Bartlett 1996). If the distance the javelin travel increases the athlete can do more mechanical work upon the javelin (Morris & Bartlett 1996). This leads to an increase in kinetic energy of the javelin (Morris & Bartlett 1996), and thereby higher release speed (Bartlett et al. 1996). The increase of the acceleration path is related to optimising the body posture. The throwing arm should be kept extended during the crossovers and at the beginning of the delivery phase (Campos et al. 2004; Morris & Bartlett 1996). Some athletes exhibit early elbow flexion before the brace leg contact, thereby decreasing the acceleration path during the crucial final period (Best et al. 1993). Again, at the end of the throw greater elbow extension increases the acceleration path distance (Chen et al. 2020).

In addition, the torso is generally leaned backwards at the beginning of the delivery phase to increase the time available to maximise acceleration (Hay 1985, Morris & Bartlett 1996 and Tidow 1996, according to Panoutsakopoulos & Kollias 2013). This backward lean with an extended throwing arm at the beginning of the delivery phase has been suggested to be an important factor for women throwers (Panoutsakopoulos & Kollias 2013). The acceleration path can also be increased by greater distance between the javelin and the right hip (right-handed throwers) which should last until the brace leg contact (Best et al. 1993). The distance between the javelin and the hip should be optimised individually in relation to one's anthropometry (Best et al. 1993). Compared to athletes with shorter throws, athletes with longer throws have demonstrated higher horizontal distance between the hip and the javelin's grip at the instant of brace leg contact (Whiting et al. 1991).

2.2 Utilisation of stretch-shortening cycle

The stretch shortening cycle (SSC) is a function of muscles where a lengthening contraction is followed by a shortening contraction (Komi & Gollhofer 1997). In their paper, Komi and Mero (1985) discussed about SSC related to the knee angle of the brace leg. At the initial brace leg contact the leg is slightly flexed and then the velocity of the flexion increases to its maximum. Komi and Mero (1985) suggested that the flexion phase should be very short and fast and be utilized in the SSC-like manner, in which the fast stretch could be used to potentiate the subsequent extension. For example, in the 1992 Olympic Games, the gold medallist Jan Zelezny had only one degree of flexion during the first half of the brace leg contact, followed by eight degrees of extension (Mero et al. 1994).

During the preparation phase the rear leg is already touching the ground. During this phase, the knee extension of rear leg can be noticed (Campos et al. 2004). From the brace leg contact up to the javelin release, the leg action patterns vary. The knee angle increases in some athletes, whereas in some it decreases (Campos et al. 2004). Panoutsakopoulos and Kollias (2013) observed that in females, the throw distance was negatively related to the rear leg knee angle at its final contact. They connected it to the possible utilisation of the SSC during the delivery phase (figure 3) (Panoutsakopoulos & Kollias 2013).

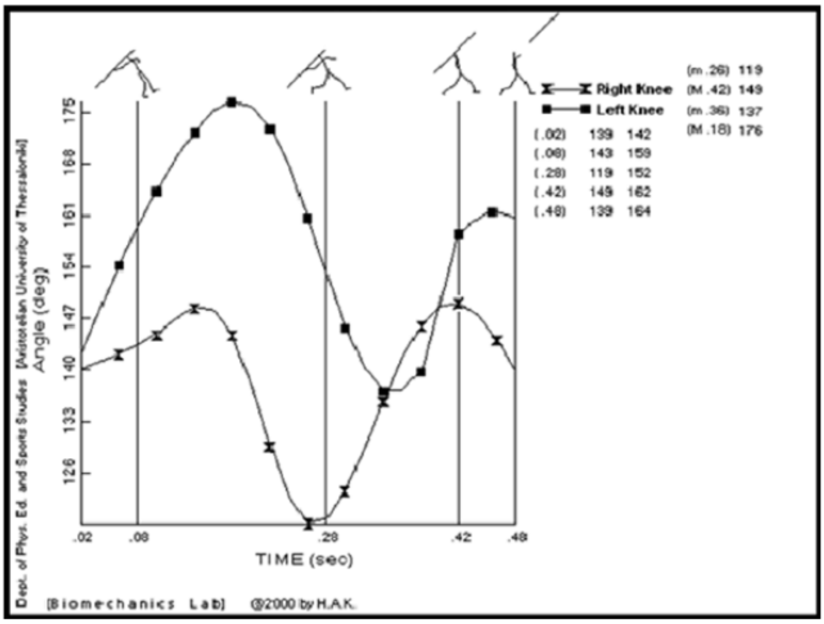

FIGURE 3. The flexion-extension patterns of both legs during the delivery phase enabling the SSC utilisation in the movement (Panoutsakopoulos & Kollias 2013).

Whiting et al. (1991) observed the knee joint angle to achieve two minima during the brace leg contact. Figure 4 shows that compared to the athletes with shorter throws, athletes with longer throws achieved both of these minima faster, and demonstrated faster extension phases, indicating a better utilisation of the SSC. Furthermore, longer throwing athletes demonstrated less knee flexion, although the final extension of the knee joint was equal for both groups (figure 4). Thus, the absolute change in final extension of the knee does not tell the whole story. (Whiting et al. 1991.) Additionally related to the SSC, during lateral trunk flexion, hip rotation, and elbow extension, many active muscles behave as an SSC action (Mero et al. 1994).

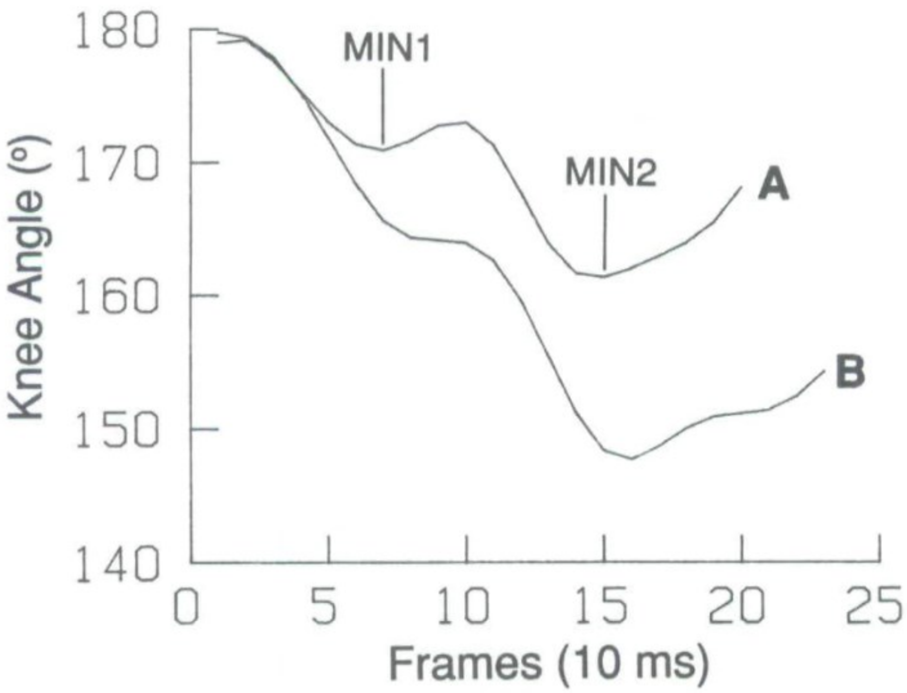

FIGURE 4. Changes in knee angle from the initial brace leg contact up to the release of the javelin (Whiting et al. 1991). A, athletes with longer throws. B, athletes with shorter throws.

2.3 The sequence of the segment motions

In addition to the SSC, the proximal-to-distal sequence of peak linear velocities of the joint centres allow for a good energy transfer through the body up to the javelin. The progressively increasing peak linear velocities of the hip, shoulder, and elbow (and in some studies wrist) during the delivery phase have been observed to demonstrate a proximal-to-distal sequence (figure 5) in many studies (Liu et al. 2014; Liu et al. 2010; Campos et al. 2004; Bartlett et al. 1996; Mero et al. 1994; Best et al. 1993; Whiting et al. 1991). A similar proximal-to-distal sequence is not equally clear perceptible in segment and joint angular motions (Liu et al. 2010, 2014).

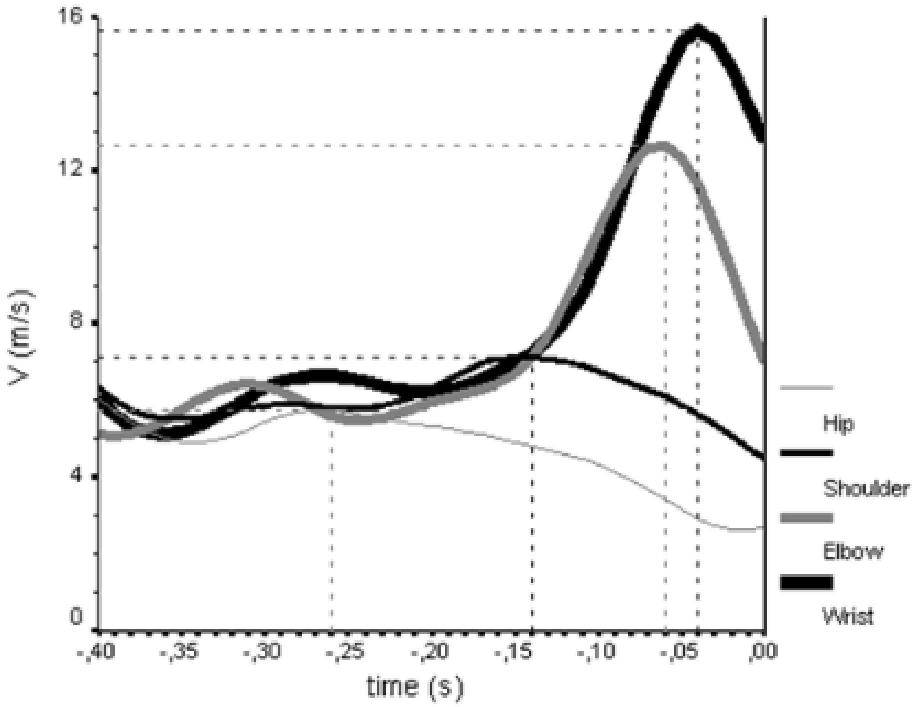

FIGURE 5. An example of a proximal-to-distal sequence of the peak linear velocities of the body segments (Panoutsakopoulos et al. 2016.) The time on the horizontal axis is in relation to the release of the javelin.

A good illustration demonstrating the transfer of momentum can be seen in figure 6. The figure above shows an 82.72 m throw and the figure below a 69.90 m throw. In the longer throw, the acceleration of the javelin is higher at the release, and a clear deceleration/acceleration curve at the elbow joint centre can be seen. Instead, in the shorter throw both the shoulder and elbow demonstrate deceleration/acceleration curves at the same time with opposite javelin acceleration. At that time, the magnitudes in changes of acceleration are smaller. Whiting et al. (1991) discussed this describing the synchronised movements of segments and transfer of momentum. The maximum elbow velocity is one of the most critical parameters related to the release velocity (Menzel 1986, according to Whiting et al. 1991), and momentum should be transferred from shoulder to elbow.

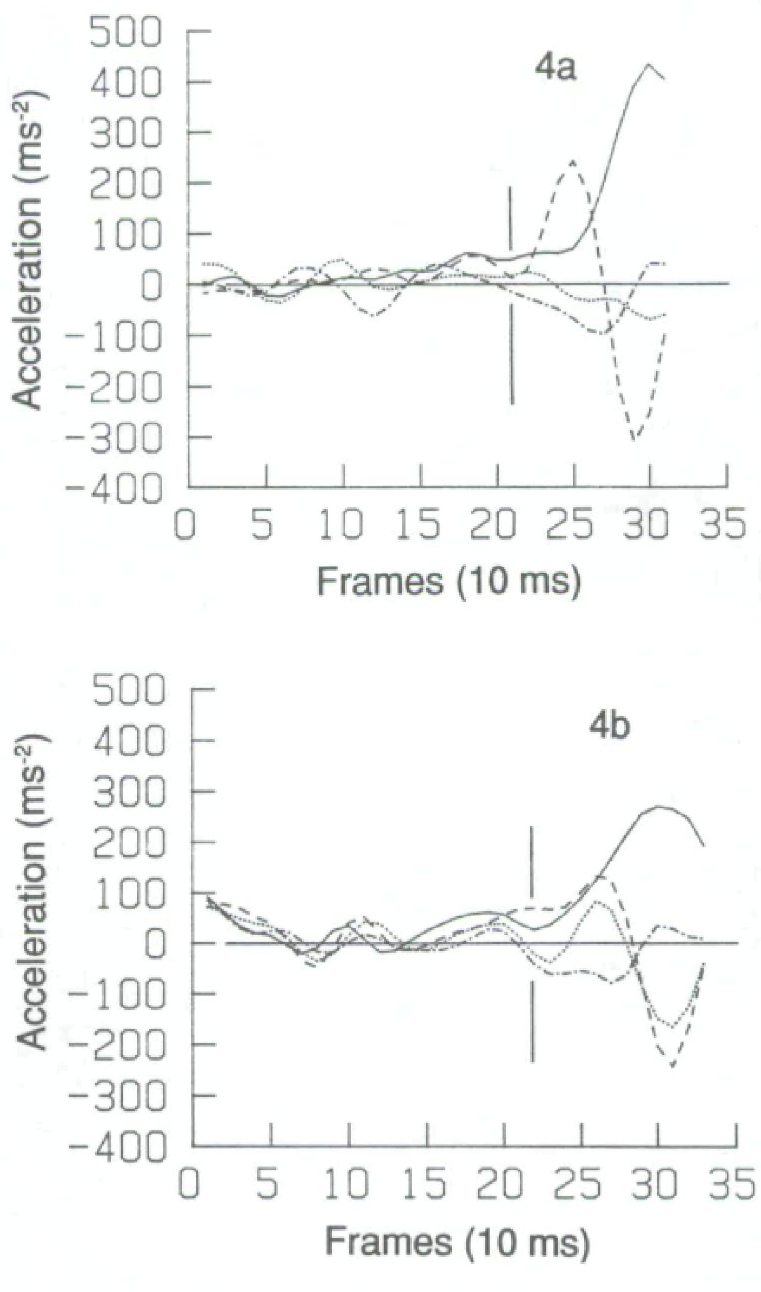

FIGURE 6. Hip (-.-.-), shoulder (….) and elbow (----) joint centre accelerations and javelin accelerations (solid line). The vertical lines represent the initial brace leg contact. (Whiting et al. 1991.)

Best et al. (1993) noticed two types of techniques related to the brace leg contact and the peak hip speed. In the first group the peak speed of the right hip relative to the centre of mass of the thrower occurred right after the final brace leg contact and this type of behaviour was found also in Campos et al. (2004) study. In the second group the peak occurred clearly before that contact (Best et al. 1993). Best et al. (1993) presented three possible reasons for these two types of phenomena: 1) there are many ways to achieve high release speeds, 2) there is one optimal technique for each thrower, but it varies between individuals and 3), the capacity to generate hip speed may be a result of other more important technique parameters, and may not be so important per se. In Best et al.'s (1993) study, one athlete exhibited both timings of peak hip speed, which supports the second suggested reason. Interestingly, the athlete demonstrated similar impulse-generating capacity (measured as the javelin release velocity minus the velocity of the athlete's centre of mass) on both types of throw. However, Best et al. (1993) suggested that generating peak hip speed too early could cause a less effective transfer of momentum through the kinematic chain and a smaller contribution from the lower body.

2.4 Individuality

When looking at joint rotations closely, it seems that there are differences between the sexes (Liu et al. 2014; Liu et al. 2010). The important observation was the difference in the timing of right hip extension (Liu et al. 2010). After the final rear leg contact, male athletes started the extension earlier than women. This suggests that men are quicker to start the hip extension, which might also allow for better maintenance of horizontal speed. (Liu et al. 2010.) Differences have also been found in the sequences of maximal angular velocities of the trunk and arm (Liu et al. 2014). Liu et al. (2014) speculated that gender differences are probably related to the differences in physical characteristics and javelin mass, as multiple computer simulations from other researchers have showed (e.g. Herring & Chapman 1992). On this account Liu et al. (2014) suggests: "The gender differences in the sequence of the trunk and arm motions should be considered in javelin throw training to improve performance and prevent injuries."

In addition to sex differences, there are observations related to each thrower having individual throwing technique (Campos et al. 2004). For example, Liu et al. (2010) found that athletes demonstrated at least two different ways to time the lower and upper segment and joint angular motions. Also Whiting et al. (1991) described individual differences in throwing performance between two throwers, suggesting a need for individualized technique analysis. Both of those throwers had throws in long and short thrown distance groups. One thrower utilised the front leg knee flexion-extension pattern suggested by Komi and Mero (1985) better and had greater hip-to-grip and pull distances, increasing the acceleration path in his longer throws. Behaviour of javelin grip movement related distances were similar for the other thrower in longer throws but otherwise different factors were responsible for longer throws. (Whiting et al. 1991.)

Mero et al. (1994) proposed that there could be different strategies to achieve the final release speed after noticing that Jan Zelezny (men's gold medallist in year's 1992 Olympic Games) had minimal angular motion in his centre of mass and the grip of the javelin compared to Seppo Räty (men's silver medallist in year's 1992 Olympic Games) who had longer contact time and demonstrated a curved pathway of the grip. Mero et al. (1994) suggested that the latter technique probably required more strength to control the centripetal force. In addition to the different strength level enabling different movement patterns, with increasing skill level, the run-up velocity could be increased to achieve higher release velocities of the javelin (Bartlett et al. 1996). If the knee flexion angle at the brace leg contact is too high, the leg cannot transfer the energy up to the upper body (Best et al. 1993). To keep the amount of knee flexion of the brace leg optimal, a good run-up velocity related to the individual's own skill level could be controlled (Bartlett et al. 1996; Bartlett & Best 1988). The throwing technique develops along with the stage of athletic development, leading to individualized technical requirements, due to physical abilities of the thrower (Kollias 1997, according to Manesh & Dr. Dhinu 2016).

To summarize, there are various factors that influence the achievement of high release speed of the javelin and long throw distance, which include certain motion patterns and the control of the knee, hip, shoulder, and elbow joints and trunk. The optimal movement pattern of the body varies making observation of kinematics worthwhile. Although, the certain basic guidelines are similar for all throwers. The utilisation of the SSC helps to achieve greater initial velocity of the javelin at the release. This means the use of flexion-extension patterns, especially in the legs. Also, the movement pattern leading to a proximal-to-distal sequence in the peak linear velocities of the body segments helps to avoid the loss of already achieved kinetic energy.

3 TRADITIONAL AND NOVEL METHODS OF MEASURING BIOMECHANICAL PARAMETERS IN JAVELIN THROWING

3D motion analysis is necessary in javelin throwing because of lateral movement in addition to sagittal movement of the thrower (Whiting et al. 1991; Best et al. 1993), although several release parameters can be estimated with 2-dimensional motion analysis (Best et al. 1993; Viitasalo et al. 2003; Panoutsakopoulos et al. 2016). In many javelin throwing related research articles the number of cameras used have been one or two (Chen et al. 2020; Panoutsakopoulos et al. 2016; Manesh & Dr. Dhinu 2016; Liu et al. 2014; Panoutsakopoulos & Kollias 2013; Liu et al. 2010; Mahmud 2009; Campos et al. 2004; Bartlett et al. 1996; Mero et al. 1994; Best et al. 1993; Whiting et al. 1991; Gregor & Pink 1985; Komi & Mero 1985).

Most research has been evaluated from competitions (e.g. Chen et al. 2020; Panoutsakopoulos et al. 2016; Liu et al. 2014; Campos et al. 2004; Whiting et al. 1991) where maximal throws can be measured. Thus, the methods used in javelin throwing research are very traditional, mostly video motion capture which has been the gold standard for throwing analysis over 3 decades (Trasolini et al. 2022). Roggio et al. (2021) describes, in recent studies, optoelectronic stereophotogrammetric multi-camera capturing system used with reflective markers attached to the body is considered to be the gold standard because of higher accuracy. However, there is only one study (Abakumova et al. 2020) that used optoelectronic motion capture to study javelin throwing kinematics. A possible reason for the lack of studies in this area could be the time-intensive nature of the method and nonsuitability to measure during competitions.

When talking about video-based motion capture which is the most used method in javelin throwing analysis, sampling frequency, digitizing errors, identification of the exact instant of release to within one frame, and data smoothing are all factors that affect the accuracy of motion analysis results (Whiting et al. 1991). With a sampling frequency of 200 Hz, the accuracy of estimates of javelin release speed is ±1–2 m/s, at 100 Hz ±2–3 m/s, and at 50 Hz ±5 m/s. (Whiting et al. 1991.) The sampling frequency also affects the ability to accurately detect the duration of the brace leg contact (Whiting et al. 1991), and higher sampling frequencies allows peak segment motion values to be captured more precisely (Chen et al. 2020). Additionally, Liu et al. (2014) reported about differences and similarities in sequences of segment motions, although they noted that their sampling frequency (60 Hz) led to timing differences that were less than the measurement error. Thus, a higher sampling frequency would have led to more accurate timing detection of segments' motions. (Liu et al. 2014.)

In addition to regular methods used in motion analysis, development of technology, such as higher computing power and graphic processing units (Galbusera et al. 2019), have led to multiple possibilities to use new ways to measure and quantify motion in sports. Elumalai and Ramakrishnan (2021) have started to develop the first real-time system for measuring javelin release parameters, height, angle, and the speed at which javelin is released, using device attached to the javelin. This system works through a cloud server and Wi-Fi. The application also saves data, allowing it to be explored later in addition to real-time use. (Elumalai & Ramakrishnan 2021.) Särkkä et al. (2016) also observed promising findings in measuring javelin throwing mechanics using a wireless device attached to the javelin in their pilot measurements, although improvements are still needed.

Technological advancements are also bringing kinematic and kinetic analysis of human via wireless wearable devices closer (Boddy et al. 2019). Wearable sensors have been used to examine a similar type of throwing motion as javelin throwing; baseball pitching motion (Boddy et al. 2019; Lapinski et al. 2019), although the results suggest that further research is needed (Buddy et al. 2019). Biomechanical modelling has still challenges in scapula (Richardson et al. 2017; Janes et al. 2012) and shoulder motion modelling. The validation of pitching motion analysis with wearable sensors can be difficult because different results in the kinematics and kinetics of pitching between two studies may be more caused by different biomechanical models used than the differences in the pitching motion itself (Gasparutto et al. 2021). Hence, the validation method used should be selected carefully according to the aims of the study.

The use of wearable devices in general type of movements such as running has been studied a lot (Norris et al. 2014). Also, artificial intelligence algorithms have been used to estimate gait related parameters from wireless wearables, but it is a relatively new approach (Caldas et al. 2017), despite the fact that the use of machine learning (ML) has been explored in biomechanics from the mid-2000s (Zago et al. 2021). One ML related motion analysis method is markerless motion analysis, which estimate human pose via teaching the ML model with labelled images to learn to recognise body parts from images it has not previously seen (Josyula 2021; Cronin

2021). Markerless systems enable measuring more people in shorter times and with lesser equipment as compared to optical motion capture systems (Roggio et al. 2021). Currently, the accuracy level of markerless approaches are more suitable for sport and rehabilitation than for clinical diagnostics (Roggio et al. 2021). Multiple open-source datasets have been released to make development and use of ML methods easier especially in field of markerless motion analysis (e.g. COCO and Human Foot Keypoint Datasets, Needham et al. 2021). In field of wireless wearables and human gait open-source datasets are not yet equal common, but those also occur (e.g. Pubudu et al. 2018). Roggio et al. (2021) proposed the future of biomechanical research being fast, fully automatic, noninvasive, and repeatable approach with a smaller number of human-dependent errors. Also, video and IMU based methods can be combined (Halilaj et al. 2021).

4 INERTIAL MEASUREMENT UNITS FOR MEASURING KINEMATICS

Commonly used wearable devices in kinematic measurements are called inertial measurement units (IMUs). In this work we focus on microelectromechanical system IMUs. IMUs are small lightweight devices which can be used almost everywhere (Roggio et al. 2021) enabling motion analysis to be done outside the laboratory (Brunetti et al. 2006). Brunetti et al. (2006) showed IMUs to be efficient in comparison to image-based methods or electrogoniometry measurements (table 1).

TABLE 1. Comparison between motion analysis technologies from Brunetti et al. (2006).

Quality	Photogrammetry	Electrogoniometry	Inertial
Cost	High	Medium	Low
Precision	High	High	Medium
CPU* Cost	High	Low	Medium
Assembly	Low	Medium	Low
Portability	Low	Medium	High
Consumption	High	Low	Low
Others	Markers occlusion	Ext. Ref.	Ext. Ref.

*CPU = Central processing unit

IMUs contain three different technologies; accelerometers, gyroscopes, and sometimes magnetometers (Godwin et al. 2009). Accelerometer measures linear acceleration, gyroscope angular motion, and magnetometer magnetic field (Lebel et al. 2017). Combination of these creates an attitude and heading reference system (AHRS) which can be used to produce 3D orientation estimation (Lebel et al. 2017). This is based on using accelerometers and magnetometers as reference systems for angular velocity measured with gyroscopes (Brunetti et al. 2006). With the gravitational force (from accelerometers) and with the magnetic pole (from magnetometers) we can get information about angular positions in global coordinate system (Brunetti et al. 2006). This is not straight-forward because signals include noise (Brunetti et al. 2006) and to be able to estimate position with IMUs, the starting point must be known (King 1998). The basic strap-down integration does not consider inaccuracies due to drift. Calibration poses and sensor fusion algorithms have been used to improve calculation of sensor orientation in relation to the global coordinate system (Robert-Lachaine et al. 2017; Caruso et al. 2021).

Sensor fusion algorithms used to reduce drift can be divided into two common main categories: Kalman filters and complementary filters (Caruso et al. 2021). For both categories, many different mathematical formulations are proposed. The orientation representation only can be presented for examples as quaternion, rotation matrix or Euler angles. The accuracy of different sensor fusion algorithms varies highly depending on the measured task. (Caruso et al. 2017.) For example, Madgwick's AHRS implementation (complementary filter) had a reported error of below one degree (Madgwick 2011), but Caruso et al. (2017) noticed in a more difficult task, human locomotion, an error above 13 degrees in Bergamini et al.'s (2014) study. Using the same implementation, measured as root mean square error (RMSE) the error in absolute angle was 4.1, 2.8, and 3.6 degrees for roll, pitch, and yaw respectively in Wada's et al. (2021) sprinting study where pelvic orientations were estimated. Sensor specificity should be also taken into account when using already optimized algorithms, as for example, Kalman filter-based trunk orientation estimation approach from Mazzà et al. (2012).

Favre et al. (2008) proposed a new method to measure 3D knee joint angle with two IMUs. They estimated orientations of the thigh and shank segments by calculating the IMU frame orientation relative to the IMU's fixed reference frame via fusion algorithm. Referring to Favre et al. (2008): "This fusion algorithm allows the tracking of IMU orientation by means of a quaternion-based time integration of the angular velocity vector as measured by the gyroscope, and by correcting the resulting orientation using the inclination estimated from the accelerometer." These orientations cannot be used to calculate joint angles before aligning IMUs to the same frame. (Favre et al. 2008.)

To calibrate IMUs to the same frame, calibration poses can be used (Robert-Lachaine et al. 2017). The accuracy can be increased with careful attachment of IMUs to body segments (Robert-Lachaine et al. 2017). The reason for careful IMU mounting location is the soft tissue artefact which could affect the accuracy of orientation and joint angle estimation (Camomilla et al. 2018). For example, the pelvis IMU should be located posteriorly because anterior placement between the superior anterior iliac spines could lead to artifacts related to abdominal breathing (Fusca et al. 2018). The location of IMU and subsequent soft-tissue artefact also affects IMU-to-segment alignment (Zabat et al. 2019). When the IMU is mounted to the segment, the IMU-axes are not yet perpendicular and same-directional with segment orientation in relation to the joint centres. The joint angles must be calculated related to anatomical orientations (Zabat et al. 2019). For example, Zabat et al. (2019) measured and calculated proposed movements related to sensor alignment to improve the accuracy of upper limb joint angle measurement.

Ferromagnetic disturbances can also cause problems in laboratory conditions because the magnetic field can vary highly (de Vries et al. 2009). Magnetic field is not constant indoors, and it can be distorted near ferromagnetic materials (Favre et al. 2008). The use of ML can enable use of IMUs without taking care about magnetometer related problems (Mundt et al. 2020c). Mundt et al. (2020c) mentioned, for example, measurements on treadmills or near railway stations which were possible despite varying magnetic fields. Even though the research focus could continue to be in improving sensor fusion algorithms (Gurchiek et al. 2019), ML-based orientation estimation of segments or straight joint angle estimation seems to be promising, especially when combined with sensor fusion. Lastly, another hinderance to IMU-based measurements could be the battery capacity and memory constraints of IMUs (Gurchiek et al. 2019). Usually, gyroscopes are the limiting units of sensors, while accelerometers can measure over 24-h continuously (Gurchiek et al. 2019).

5 BASICS OF MACHINE LEARNING

Machine learning (ML) is a subcategory of artificial intelligence that enables computers to detect patterns in data (Alzubi et al. 2018). While in general programming a human write rules to get answers based on data, in ML data and answers produce rules by which answers can be obtained from new similar data (figure 7) (Chollet 2021, p. 4). ML systems are more trained than programmed (Chollet 2021, p. 4). Chollet's determination refers mostly to supervised learning. ML can be divided into three subgroups: supervised learning, unsupervised learning, and reinforcement learning (Wuest et al. 2016). In supervised learning both the input features (continuous, categorical, or binary) and corresponding output values are input to the model (Kotsiantis 2007). Instead, in unsupervised learning, the outputs are not given to the model, but the model divides the data into clusters. In reinforcement learning the learner discovers a certain environment and tries different options and changes it functions based on feedback from the environment. (Kotsiantis 2007.) Supervised learning is divided into two subgroups: to classification (table 2) and regression algorithms (Alloghani et al. 2020, p. 4).

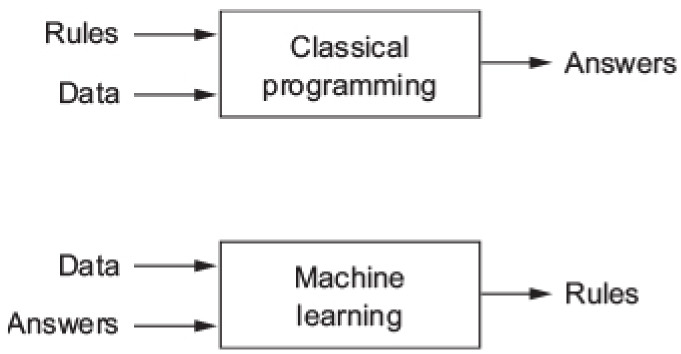

FIGURE 7. Chollet's (2021, p. 4) illustration of differences between programming and ML.

TABLE 2. An example of data format in supervised classification (Kotsiantis 2007).

Data in standard format					
Case	Feature 1	Feature 2	…	Feature n	Class
1	xxx	x		xx	good
2	xxx	x		xx	bad
3	xxx	x		xx	bad
…					…

Supervised learning is commonly used in biomechanical approaches. There are many specific supervised learning approaches where neural networks (NN) are the most used (Gurchiek et al. 2019). A simple neural network is called a perceptron (Krogh 2008). NN can be divided into three categories: artificial neural networks (ANN), convolutional neural networks (CNN), and recurrent neural networks (RNN). In some sources ANN refers to all types of neural networks or in general networks including multilayer perceptron. In some sources ANN only refers to feedforward neural networks (FFNN), in which case the division to ANN, CNN, and RNN is used. If using multiple ANN layers to process complex data, the term deep learning is used (Josyula 2021). The difficult part of supervised learning is the danger of overfitting or underfitting (Galbusera et al. 2019), which can happen if the learning model is not complex enough to detect features of input data (Galbusera et al. 2019). This problem can often be overcome by using larger, more variable training datasets.

The classic phrase “garbage in, garbage out” is important in fields applying ML (Geiger et al. 2021). The input data in supervised ML systems is “ground truth” from which the model learns. If the input is problematic the output is also problematic. Thus, the input data determines the quality of the supervised model. Geiger et al. (2021) reviewed how input data has been labelled in life- and biomedical sciences and how valid the labelling methods are. They found that serious consequences can occur if the labelling is not done accurately. One example described by the authors is related to detection of criminals from face images. The images were labelled from prison mug shots and noncriminals from professional social network profiles (Wu & Zhang 2017). A crucial phenomenon that the model learned was that criminals did not smile but noncriminals smiled and the output of the model was biased accordingly (Bergstrom & West 2020, according to Geiger et al. 2021). (Geiger et al. 2021.) Here the labelling refers mostly to the image labelling which can be used in biomechanical fields in many ways, for example in markerless motion analysis. The same human-labelling validity phenomenon can also be seen in all other measurement areas: selecting the right type of modelling methods to get the right type of “ground truth” for a new supervised model.

The importance of human-labelling accuracy depends on the field the information is used in. Labelling MRI images as cancerous or not cancerous is more important to label correctly than classifying product reviews as positive or negative (Geiger et al. 2021). The method used to measure or label the input data for ML models should be at least as accurate as the result must be. In kinematic measurements and in javelin throwing the validity of the input data depends on the “ground truth” values used to determine some important technical parameters from performance.

6 KINEMATICS VIA MACHINE LEARNING AND IMUS

In a systematic review from 2019 (Gurchiek et al. 2019) only 9 studies used IMUs (inertial measurement units) to estimate kinematics or kinetics. After 2019 many studies have used greater number of participants and the number of studies using IMUs and ML (machine learning) has increased significantly (e.g. Tan et al. 2022; Sung et al. 2021; Sharifi Renani et al. 2021; Senanayake et al. 2021; Rapp et al. 2021; Mundt et al. 2021; Hernandez et al. 2021). Most ML related studies estimating kinematics and kinetics until 2019 were subject-specific models (80%) (Gurchiek et al. 2019). The tasks where the models are used to predict joint angles are mostly walking and running (e.g. Sung et al. 2021; Senanayake et al. 2021; Hernandez et al. 2021). Few studies have predicted lower limb angles during functional activities (Tan et al. 2022) or lower limb exercises (Argent et al. 2019). This chapter present prediction of kinematics with ML and IMUs related to four mentioned tasks, Tan et al. (2022) and Argent et al. (2019) studies being only ones not being walking or running studies.

In biomechanical approaches many different types of machine learning (ML) have been utilized in joint angle estimation of IMU data. The deep neural networks (NN) (more than three layers) seem to be most common on the kinematic analysis of human motion area (Gurchiek et al. 2019). In this thesis the model types used to estimate kinematics in literature were simpler artificial neural networks (ANN) (Mundt et al. 2021; Mundt et al. 2020b; Mundt et al. 2020c; Lim et al. 2019), long short-term memory recurrent neural networks (LSTM) (Mundt et al. 2021; Rapp et al. 2021; Sung et al. 2021; Liang et al. 2021; Sharifi Renani et al. 2021; Mundt et al. 2020a; Mundt et al. 2020c), convolutional neural networks (CNN) (Mundt et al. 2021; Chow et al. 2021; Rapp et al. 2021; Dorschky et al. 2020), and a generative adversarial network (Senanayake et al. 2021). Hernandez et al. (2021) used a combined model of the CNN and LSTM. Other machine learning methods used for estimating kinematics are linear regression, polynomial regression, decision tree regression, and random forest regression (Argent et al. 2019). Also, support vector machine algorithms have been used for estimating linear accelerations and angular velocities of the body segments (Carter et al. 2022).

The difference between the standard fully connected feedforward neural networks (FFNN) and the recurrent neural network (RNN) is an additional loop in RNN (figure 8). The LSTM is one type of RNN, and it contains gate layers (Mundt et al. 2020a). Mundt et al. (2020a) describes how the gate layers ‘remember’ the information from the long phase and the short time relevant information can be discarded. LSTM have combined with FFNN or CNN in few studies (Tan et al. 2022; Sharifi Renani et al. 2021; Hernandez et al. 2021). The combination of FFNN and LSTM is called as BiLSTM (Tan et al. 2022). Human movements are continuous which could provide the success with using RNN in estimation of body kinematics referring to the time-dependency.

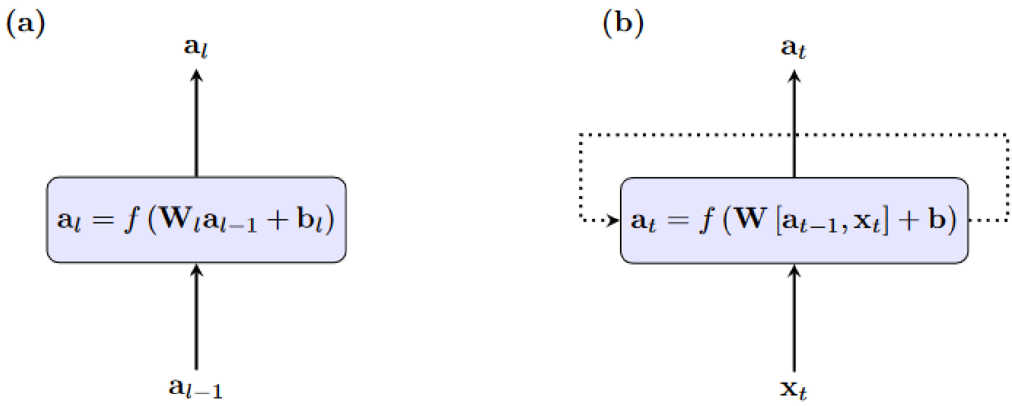

FIGURE 8. Structure of the standard fully connected feedforward neural network (a) and a recurrent neural network (b) (Mundt et al. 2020a).

According to Gurchiek et al. (2019), study from Xia et al. (2018) were first using the combined CNN and LSTM layers in kinematics and kinetic research area, although Xia et al. (2018) used surface electromyography sensors as an input. Until 2019 models including purely LSTM layer or CNN and LSTM layers together were not used to predict joint angles in studies accepted by Gurchiek et al. (2019). To my knowledge, Mundt et al. (2020a; 2020c) were the first to use LSTM to predict joint angles. Mundt et al. (2020a, 2020b, 2020c) studies are not accepted with Gurchiek et al. (2019) criteria because they used synthetic IMU data which was derived from marker data. LSTM and CNN-LSTM models with real IMU data have used starting at the latest since 2021 to predict joint angles (Tan et al. 2022; Hernandez et al. 2021; Liang et al. 2021; Sharifi Renani et al. 2021; Sung et al. 2021).

Rapp et al. (2021) mentioned that the RNN is better for offline processing comparing to CNN. Mundt et al. (2021) noted that LSTM requires less data pre-processing which makes prediction faster and easier in real-time applications. Mundt et al. (2021) seems to be only ones to have compared different types of NNs. Their result was that the LSTM had clearly higher (worse) final loss than other NNs. The pretrained CNN was not better than the multilayer perceptron or new CNN. When looking accuracies more closely the LSTM had better estimation ability for hip extension-flexion and adduction-abduction, and knee flexion-extension angles than the multilayer perceptron. For hip internal-external rotation, knee adduction-abduction and internal-external rotation, and for all ankle joint angles the LSTM was worse than the multilayer perceptron. The CNN was more accurate than the LSTM or the multilayer perceptron for all angles. Again, while using the leave-one-subject-out validation the LSTM performed generally worst compared to other tested models. Time normalization was the discussed reason for the results because joint angle data have lot of variation in initial values at foot contact and this behaviour cannot be captured well with multilayer perceptron and LSTM. Regardless of the differences in model accuracies all the models gave very good overall results. Even though the research gave a general overview to the use of different types of models, the authors did not give recommendation to use a particular model. (Mundt et al. 2021.)

Information of the devices can be used as raw data or in calculated format to certain efficiency variables. Lim et al. (2019) used 2-dimensional acceleration information from the sagittal plane to calculate estimated behaviour of centre of mass (including the drift remove to the calculations) and used this information (acceleration, velocity, displacement, and time) as an input for the ANN model. The Madgwick orientation algorithm has been utilised also to create input for machine learning models (Argent et al. 2019). Orientation-based machine learning models performed better in many tasks than raw data models in predicting different lower limb exercises (Argent et al. 2019). Raw data seemed to be good when predicting knee extension, hip flexion, or straight leg raise (Argent et al. 2019). Using raw accelerometer and gyroscope data could be a good way to avoid using magnetometers in environments where the magnetic field varies a lot in short distances (Tan et al. 2022). According to Tan et al.'s (2022) discussion the ML models using real IMU data are not yet as accurate as Kalman filter-based approaches.

The estimation of sagittal joint angles is perceived to have generally good accuracy (Mundt et al. 2020a). Mundt et al. (2020c) observed that transverse plane predictions are more difficult compared to predictions in sagittal plane via both FFNN and LSTM models. Hernandez et al. (2021) noticed same phenomenon with combined model of CNN and LSTM while examining accuracy of prediction of lumbar rotation. Hernandez et al. (2021) discussed this being probably caused by higher marker placement errors based on Osis et al. (2014) research. Also, according to Mundt et al. (2020a) the inaccuracy of the transversal or frontal plane angles may be caused by larger variances in these angles (Mundt et al. 2020a). The IMUs can detect very small motions and if the order of the inputs is small the variation of the angles affect more than in greater angles as a sagittal knee angle. Increasing the size of the training dataset and variance in data would help address this phenomenon. (Mundt et al. 2020a.)

The prediction of kinematics with ML and and IMUs requires huge amount of data. Rapp et al. (2021) observed that model trained with hundreds of participants improved model accuracy compared using tens of participants. Rapp et al. (2021) used simulated IMU data to study this phenomenon (figure 9). Because there is requirement for large corpora of data also other researchers have used marker-based motion capture data to generate simulated IMU data to estimate kinematics (Sharifi Renani et al. 2021; Rapp et al. 2021; Mundt et al. 2021; Mundt et al. 2020a; Mundt et al. 2020b; Mundt et al. 2020c). Mund et al. (2020c) remarked that soft tissue movements are not included to the simulated IMU data and that could be taken into consideration. Mundt et al. (2021) used 23 participant's ground truth IMU data to create 23 different variations of other participants data. Based on Mundt's et al. earlier studies the use of simulated IMU data in the NN models can overestimate the accuracy of the predictions but still be a good way to improve the base of the models in addition to the real IMU data (Mundt et al. 2021). Mundt et al. (2020b) noticed that simulated data improves prediction of kinematics but not kinetics.

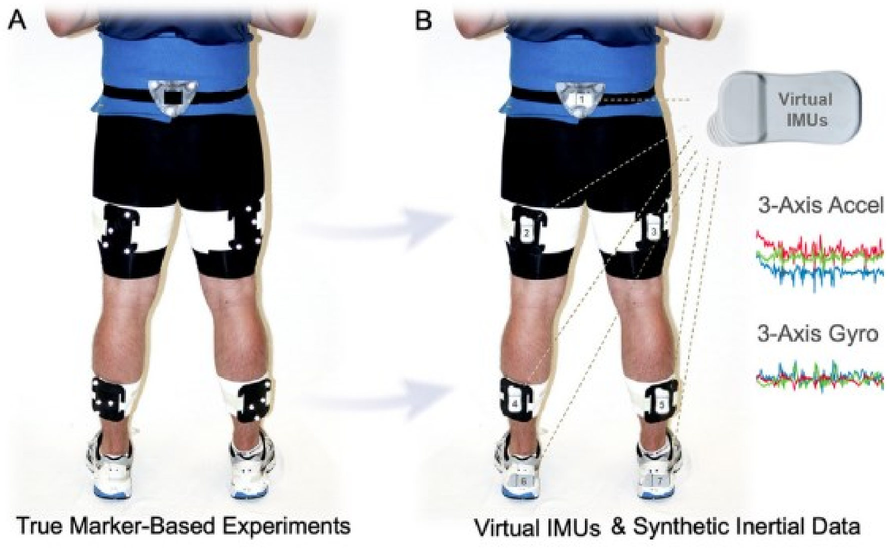

FIGURE 9. Virtual IMUs were created using clusters of markers (Rapp et al. 2021).

Feature extraction is an important part of building the model by giving a better accuracy for the estimations selecting for example only certain IMU signals to the ML model (figure 10) (Sung et al. 2021). Also, Dorscky et al. (2020) selected signal inputs using only sagittal plane signals. They also created their own CNN for each output variable. Magnetometer data often interfere with prediction accuracy of deep learning models (Rapp et a. 2021). Anthropometric measurements can also be added to the model as input variables. Mundt et al. (2020c) used participants' height, weight, width of pelvis, and length of foot, thigh, and shank to the final models. Argent et al. (2019) tested height, weight, age, sex, and limb segment lengths.

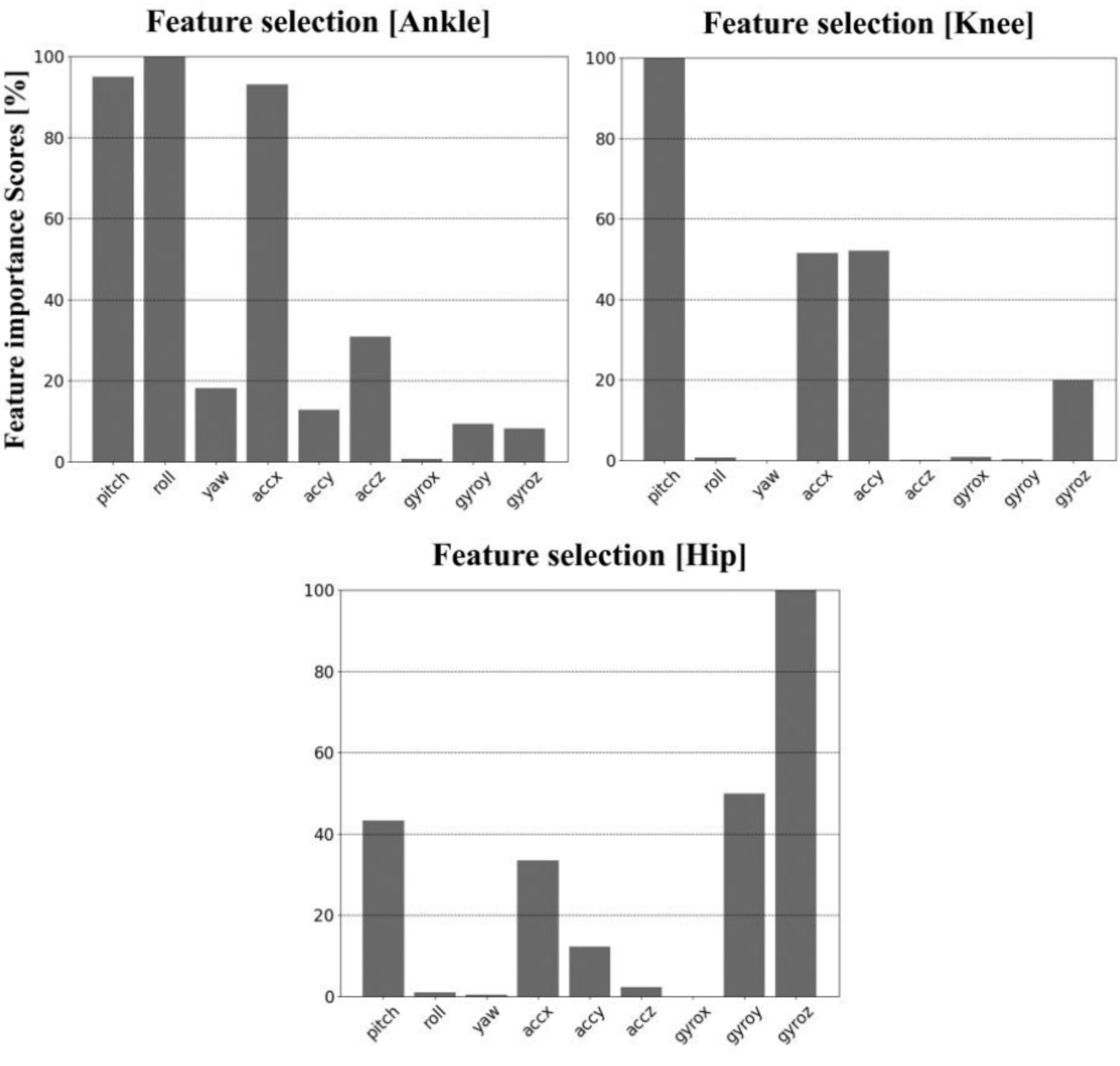

FIGURE 10. An example of feature selection (Sung et al. 2021). For hip angle prediction the y- and z-axis gyroscope signals were significantly more important than for ankle or knee angle prediction.

Mundt et al. (2020a) focused on comparing different feature selections for LSTM model. With different feature selection methods, they noticed that using three sensors led to higher accuracy than the use of five sensors. The accuracy of the sensors located on the pelvis and shanks were better than the location on the pelvis and thighs. This is because shank sensors measure motion of the whole leg whereas thigh sensors measure only the motion of the upper leg. The pelvis sensor can be described as a measure of the centre of mass. If the measurement does not require high accuracy the use of only a pelvis sensor may be sufficient. (Mundt et al. 2020a.) Tan et al. (2022) compared singe- and double-leg models. They noticed a task-dependency in which one of the models worked better. In total the single-leg model was more accurate, but the double-leg model was better if activity required reciprocal movements of both legs, for example walking. (Tan et al. 2022.)

7 RESEARCH QUESTIONS

The inspiration for this thesis came from the study of Hernandez et al. (2021), in which they used five IMUs to estimate lower limb kinematics via a combined CNN and LSTM model. Their goal was to avoid effects of drift of IMUs, and calibration poses. The model was designed to learn the variation of IMU placement in different subjects. In this work the aim is to examine whether machine learning models and IMUs could be used to estimate lower limb kinematics in javelin throwing. The interesting time moments to examine certain parameters were selected based on literature and recommendations from a biomechanist from the Research Institute for Olympic Sports with whom the research was conducted.

Question 1: How precise pelvis tilt, pelvis list, pelvis rotation, and hip and knee flexion of rear and front legs (defined based on final contacts) estimates can be obtained from the last crossover stride until javelin release via 6-axis IMU-based individual LSTM model or general LSTM group model with a small amount of data (ten participants)? As a secondary question, which one of the models is the most accurate?

Hypothesis 1: The lack of training data is a general reason for difficulties in machine learning model development (e.g. Sharifi Renani et al. 2021; Rapp et al. 2021), which is possibly why with a small participant group (ten), the general model does not make accurate predictions; error below five degrees. Individual models are probably more accurate than the group model, but there may be bigger problems related to robustness of individual models. The variation of the data can be small so the model learns well but it is overfitted and cannot predict future data from other measurements or participants if throwing the pattern changes even slightly.

Question 2: Can the accuracy of the prediction of the front leg knee angle (the supportive leg in the last phase of the javelin throw) from last crossover stride until javelin release be increased compared to a general group model using only joint specific IMUs (on thigh and shank segments) and this knee angle being the only target variable of the LSTM model?

Hypothesis 2: The hypotheses for this question are not clear. For example, Favre et al. (2008) used sensor fusion and only two IMUs on the thigh and shank segments to estimate knee angle. In sensor fusion algorithms more sensors are not needed to estimate a single joint angle. In some studies (e.g. Mundt et al. 2020a) reduced number of sensors has led to more accurate predictions. Extra sensors may lead to bias in knee joint angle estimation, although Tan et al. (2022) found that the double-leg model worked better if the task required the use of both legs. This could mean that additional sensors help the LSTM model to learn important features.

8 METHODS

In this work, different variations of the long short-term memory (LSTM) layers were compared to the gold standard method: 16-camera Vicon-system and inverse kinematics (IK) method performed using Opensim software. All LSTM models were trained using IMU data. The input data was 3-axis accelerometric and 3-axis gyroscopic data. The joint angles derived from the gold standard method were used for supervising learning. The models were built just for observing the possibilities of this kind of method, and were not meant to generalize to a larger population or even for those athletes whose throwing data were used to train the model. Different models were built based using data from all participants in one model or individual models separately for individual participants. Test sets of the models were separated from training data using modified leave-one-subject-out (group models) and leave-one-throw-out methods (individual models).

The predicted variables were pelvic tilt, pelvic list, pelvic rotation, right hip flexion, right knee flexion, left hip flexion, and left knee flexion (figure 11). For all participants, the left leg was the front leg, and the right leg was the rear leg, based on the final feet contacts of javelin throw performance before the release of javelin. Later the rear and front leg terms are used. All pelvic orientation and hip and knee flexion values are presented in relation to standing position. In pelvic orientations positive directions denote rotation to the left, backward tilt and list as the right pelvis being lower than the left. The absolute orientations and joint angles were prediction targets because they allow angular velocities to be derived if desired. In this work, linear velocities of the joint centres were not processed.

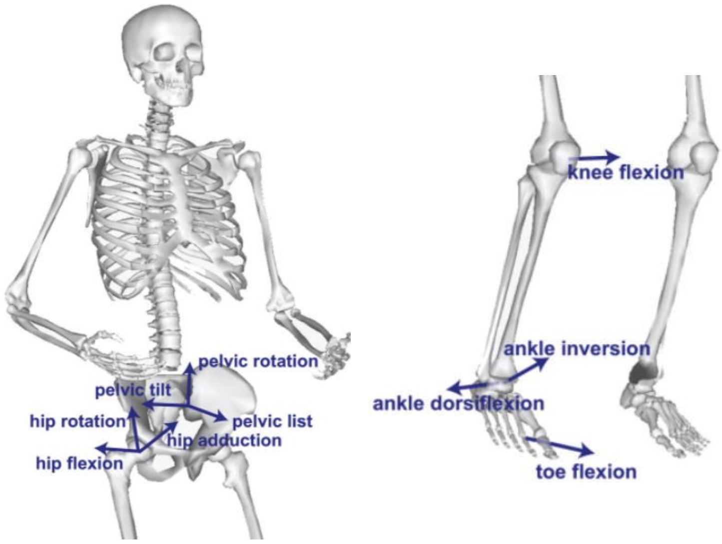

FIGURE 11. Pelvic orientations and hip and knee joint angle presentations in Opensim. (Image: https://europepmc.org/article/pmc/pmc5507211)

8.1 Participants

Inclusion criteria for participants were (1) 15–40 years old, (2) purposefully practising javelin thrower, (3) able to do maximal javelin throws, (4) can understand information in Finnish or in English provided that the research group was able to ensure she/he understood all aspect of the research to be able to provide informed consent. Exclusion criteria were current musculoskeletal injuries.

Participants were recruited by KIHU (Research Institute for Olympic Sports in Finland) after the ethical approval. Ten purposefully practising javelin thrower athletes participated in this study (seven men, three women). Their average age (± sd) was 23.6 ± 3.4 years. Participation to the study was voluntary and everyone had to right to refuse to participate. Participants had rights to get information about what participating mean and how the collected data will be used and stored. Information of research was told in literary form in paper and in electrically. Information was also told orally before starting the testing session. At the beginning of the testing session information about voluntary aspect was iterated. Participants had a possibility to suspend their participation at any time without giving a reason. Participation to the research included one testing session. In addition to other forms, participants also signed the agreement for the release of images. Measurements were carried out at Kuortane Olympic training center in an indoor track and field hall. Throwing performances were performed towards a javelin throwing tarp. The study protocol was reviewed by the University of Jyväskylä Human Research Ethical Review Board, and approval was granted on February 24, 2021 (170/13.00.04.00/2021).

8.2 Devices

A 16-camera (Vero v2.2) Vicon system (Oxford Metrics, Inc., Oxford, UK) was used for the marker-based motion capture (sampling frequency 300 Hz). A 45-marker set (Vicon full body Plug-In gate and extra markers on medial side of elbows, knees, and ankles) was used. Vicon IMeasureU BlueTrident sensors (IMeasureU, Auckland, New Zealand) were connected to the Vicon system via an external Bluetooth 5.0 dongle. The total number of sensors attached to the participant was 11, but ultimately 5 sensors (at the pelvis, thighs, and shanks) were used for models in this thesis. For the IMU sensors, the sampling rate was 1200 Hz in the Nexus 2.11

software (Vicon Motion Systems, Ltd., Oxford, UK) (the actual sampling rate of accelerometers and gyroscopes was 1125 Hz, magnetometer data was not used).

Calibration of the devices. For the Vicon system, cameras were first zoomed and focused on calibration volume. For all cameras additional settings in Nexus software were good keeping those as default. Unwanted reflections were masked out. Cameras were calibrated using T-shaped 5-marker wand moving it around calibration volume to get at least 1000 frames of valid wand data (default number of valid frames was used). After calibration, the volume origin was set. Calibration volume of the Vicon system was approximately 8m x 3m x 3m. The cameras were set up around the javelin throwing track in indoor hall (figure 12). Only the raw data of IMUs were collected, so the calibration of the IMUs was not required.

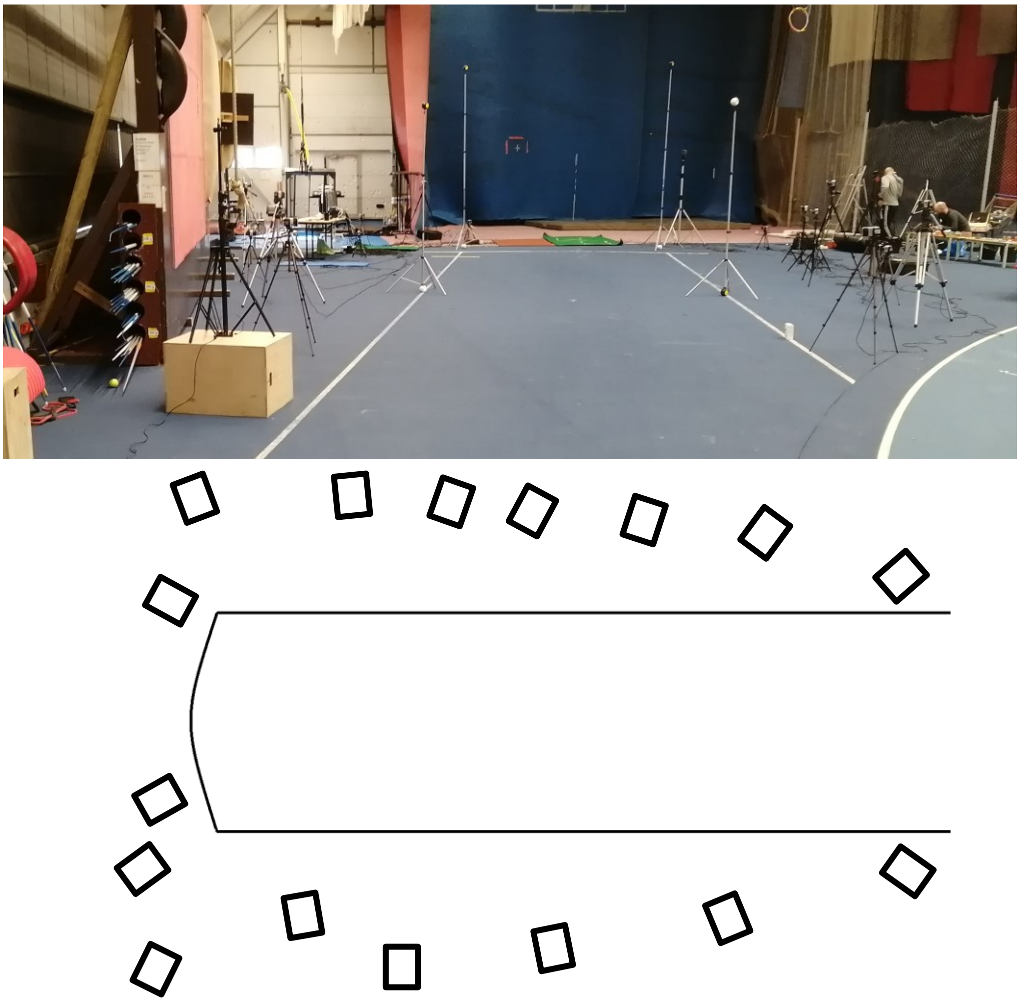

FIGURE 12. Above image of the throwing area at Kuortane Olympic training center. Below estimated positions of the Vicon-cameras around the throwing area. The cameras were tried to set to capture the area near the foul line the most accurately because the velocities of the motion are fastest there.

Placements of the sensors. All sensors (figure 13) were mounted to segments such that soft tissue artefact was minimised: on the shanks on the tibias, on the thighs in mid or slightly distally on the lateral side, and on the pelvis between the posterior superior iliac spines (figure 14). Placements of the sensors were not measured, only estimated roughly to get phenomenon about the sensor placement variety. Chow et al. (2021) says that the attachment of the IMU at the anteromedial tibia being more convenient for users. This supports the idea the athletes could attach the sensors in these placements by themselves in field use. The shank and thigh IMUs were positioned under the support bandage after applying double-sided tape. The comfort of the tightness of support bandage was asked from participants. If athletes felt the support bandage uncomfortable or may affect their performance, the backing was prepared being looser. The pelvic sensor was attached with the double-sided tape and hospital tape (Leukoplast). For practical reasons, the pelvis sensor was attached for some athletes on the skin and for some on the belt.

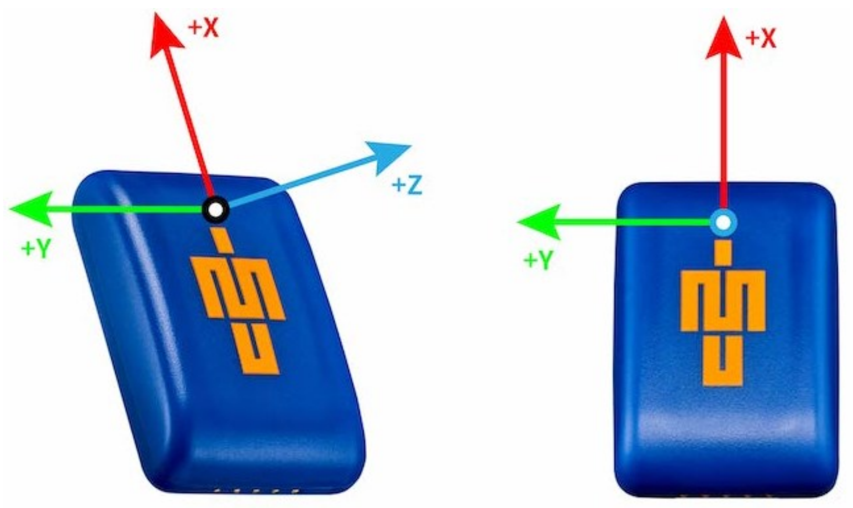

FIGURE 13. The axes of BlueTrident IMUs. (https://logemas.com/knowledge-base/faq/vicon-blue-trident-axes/)

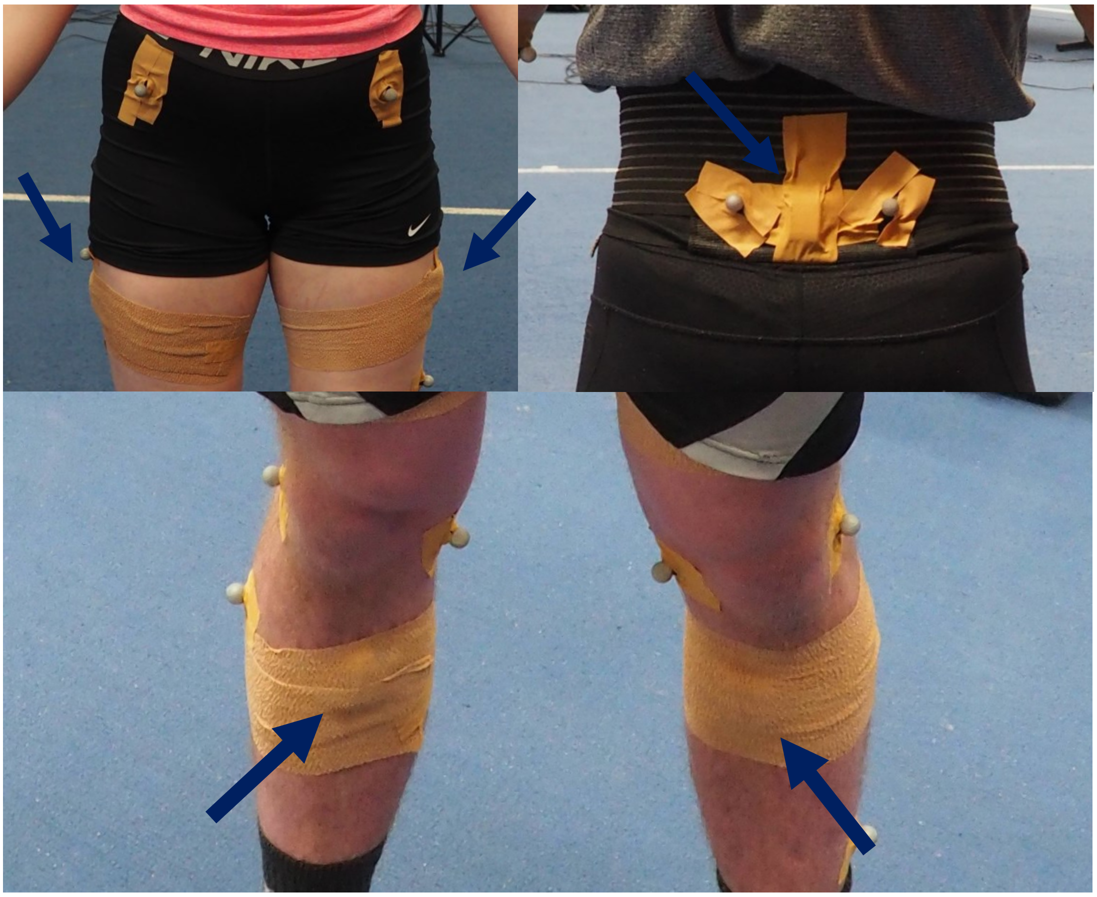

FIGURE 14. The placement of the thigh, shank, and pelvis IMUs (IMU locations shown with blue arrows). All IMUs were located x-axis vertically (upwards positive), and y-axis to the left on the surface.

8.3 Protocol

Participants were instructed to do their own ordinary warmup just before measurements and informed about markers and sensors related preparation time which may lead to need for other short warmup before actual measurements. The duration and contents of the warmup between participants varied. When athlete came to measurement area the reflective markers for optoelectronic system and inertial measurement units were mounted to body. This took approximately 30-50 minutes. Because the long preparation time, athletes got opportunity to continue their own warmup for maximum 20 minutes before data collection. Most usually this warmup included short runups and throwing motion implementations. After second warmup subject calibration trials were captured to enable subject specific automated marker labelling in Nexus. At the beginning of that trial athlete stood in “motorcycle” position approximately 5 seconds and after that participant was instructed to move her/him body to find range of motions for different body segment motions to help Nexus automated labelling to find markers in javelin throwing trials. Data collection was started with low intensity throws and intensity increased steadily. Finally, from six to ten maximal intensity throws were performed if possible. Any pain or uncomfortable feeling led to cessation of measurements. At the end, markers and sensors were removed. The duration of measurement session including warmups was approximately three hours.

8.4 Data pre-processing

Marker data from Vicon was first analysed with the Nexus software (v2.11). First, subject specific calibration models for each athlete were created. Additional markers were used in medial side of knees and elbows. Thus, individual models were created utilising Vicon's Plug-in Gait full body model. Recorded data was digitised with automated labelling and visually verified using Nexus. 3D coordinates from all labelled landmarks were exported as C3D files. After that, Matlab (R2020b) was utilised with C3D converter script to get TRC files which can be read via OpenSim. Before exporting the TRC files, gaps in the marker trajectories were filled using linear interpolation.

OpenSim 4.3 (SimTK) was used to calculate IK. To calculate IK, the Gait2392 model (Delp et al. 1990) was used. The model was scaled individually for each thrower based on the body mass preserving mass distribution. In IK settings the weight for markers can be given. That affects how important a marker is for each OpenSim model and how much error there can be. In this research, all markers were set to value 1, the default value. From OpenSim, kinematic data was exported as STO files.

8.5 Building deep learning models via Python3

A LSTM (Hochreiter & Schmidhuber 1997) based neural network was built to predict hip and knee joint angles as time series (Sung et al. 2021; Sharifi Renani et al. 2021; Rapp et al. 2021; Mundt et al. 2021; Liang et al. 2021; Mundt et al. 2020c; Mundt et al. 2020a) and pelvic orientation angles. The pre-processing of the data continued with Python 3 in the Jupyter Notebook environment (Jupyter Team, https://jupyter.org). The phases from the last crossover stride to release of javelin (approximately) from both submaximal and maximal throws were used. Total amount of the throwing trials was 167. Parallel with pre-processing data, the joint angles, 3D acceleration and angular velocity data was checked. Outliers were double checked and the reason for data loss or unusual shape was attempted to be identified. Few times, the reason for weird looking data related to all data set was probably related to bad synchronisation of IMUs with Vicon system or system breakdown because the lower sampling frequency was found from data logs. Some of the throws were excluded based on research minutes if there was a mention about loosening of a sensor. If the reason for outliers could not be found the throw was not removed. The final number of used throws was 141.

8.6 Filtering and selection of the phase

The joint angle data was filtered using fourth order low-pass Butterworth filter with the cut-off frequency of 32 Hz because fast movements in javelin throwing performance. The low pass cut-off value defines the higher frequency signal cut out. Thus, the use of 32 Hz cut-off with joint angles instead of general 8 Hz cut-off used for marker trajectories is a better option in order avoid data loss. IMU data was filtered with fourth order low pass Butterworth filter using the cut-off frequency of 90 Hz. As in the previous case, a high cut-off value was used to avoid data loss. For example, Hernandez et al. (2021) even used raw IMU data in their CNN-LSTM model. Post filtering, IMU data was down sampled to the same sampling frequency than the marker data (300 Hz). Joint angle data was not up sampled to 1200 Hz for reasons related to practicality.

After the separate pre-processing of signals, joint angle data and IMU data were combined, and data was trimmed from the beginning of the last crossover stride to the release of the javelin. The beginning of the last crossover stride was determined with graphical view of the marker data as the instance when the height of the left toe marker started increasing. The release moment of the javelin was determined via videos (extra video cameras from KIHU, calibrated via LED light connected to Vicon system) as a time moment of the frame when the javelin is touching the throwing hand just before release. At the end, all throws were cut to a 184 points timeseries. The length of timeseries is based on shortest accepted timeseries length. Thus, from few throws the absolute javelin release was cut out. The aim was to predict whole timeseries and the exact time of javelin release was not important. The MinMaxScaler from Python sklearn library (Scikit-learn, Pedregosa et al. 2011) in range from -1 to 1 was used to normalize data.

Both input and target parameters were scaled. Finally, the data were shaped to the form which was accepted for the LSTM layers.

Group model and individual models with all five sensors predicting all target parameters (pelvic orientations, hip, and knee flexions), and group model with the thigh and shank sensors of front leg predicting front leg knee flexion were tuned and trained using both maximal and submaximal throws but only maximal throws were used as test data to select hyperparameters and to test models at the end. This study was a proof of concept and therefore illustration about phenomena of models' learning. Therefore, individual models were built only for five subjects having over 15 throws. These five subjects had from eight to thirteen submaximal throws and from five to ten maximal throws.

8.7 Hyperparameter tuning

The development and training of the models were implemented using TensorFlow 2.7 machine learning platform (Abadi et al. 2016) and Keras 2.7 deep learning API (Chollet et al. 2015). The numbers of units were tested between 32–1024 using step size 32. Tested dropout values were between 0.1 and 0.8. Other hyperparameters of LSTM layers were not tuned. That means, the default activation function was tanh and recurrent activation function was sigmoid. To compile the model, the commonly used ADAM optimizer (Kingma et al. 2015) was used but the learning rate for it was tested. The loss and accuracy metrics were collected as mean squared error (MSE). The maximum epochs in tuner were set to 10 to avoid long tuning times. This reduces the possibility to find optimal parameters. In addition, the patience value five was used for early stopping.

At the end, the units and dropout used in LSTM layers and learning rate for model optimizer were selected checking which of hyperparameters are the most popular through all different training sets (from leave-one-subject or leave-one-throw out implementations). Thus, all data was used for selecting hyperparameters. The selection is based on the hyperparameters which gave best predictions. It could also be that in the same tuning round, there were found larger errors. This selection method saved training time. After finding best hyperparameters, the number of optimal epochs was also searched. This was done by training the model with optimal parameters using all training data for certain group model (not included testing data). Maximum number of epochs tested was 500. The selected hyperparameters for final group models are in table 3.

TABLE 3. Modified hyperparameters in training. Individual hidden sizes, learning rates and number of epochs for five individual models.

Hyperparameters	Group	Group-Knee	Individual
Number of layers	2	4	2
Batch size	16	16	16
Hidden size	320	320	416, 528, 352, 192, 320
Optimizer	Adam	Adam	Adam
Learning rate	0.003	0.003	0.003, 0.003, 0.005, 0.005, 0.005
Number of epochs	100	100	400, 400, 280, 450, 430
Activation	Tanh	Tanh	Tanh
Recurrent	Sigmoid	Sigmoid	Sigmoid

TimeDistributed layer and the used activation function effect was tested before real hyperparameter tuning with one smaller data group. This layer is generally used to shape data to forecast future values. This or Dense layers did not seem to increase the prediction accuracy visually. Hence, these were not used. Also, the opportunity use only one LSTM layer was tested, which reduced data smoothness. In addition to ADAM optimizer, RMSprop was tested. Both seemed to be better for some variables. At this point, the general ADAM optimizer was selected. The implementation of the model followed the sequence illustrated in figure (15) other than cross-validation was replaced with modified leave-one out methods.

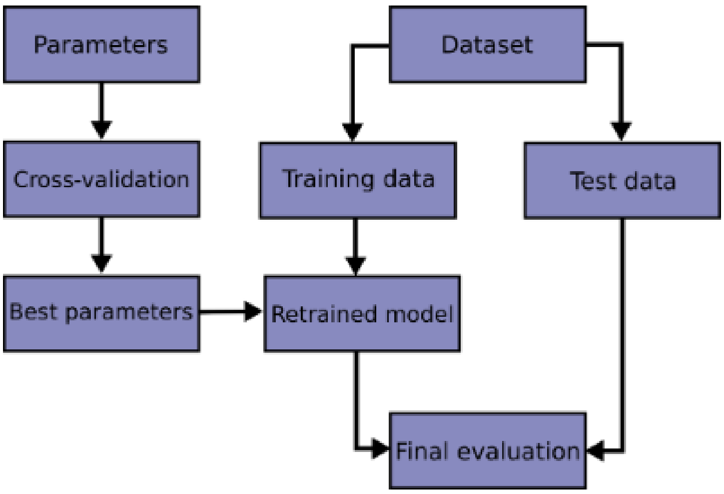

FIGURE 15. Used sequence in training the models. (Graph: https://scikit-learn.org/stable/modules/cross_validation.html)

8.8 Statistical analysis

Before the statistical analysis, the joint angle data was converted to absolute degree values using same MinMaxScaler which was used before training the model. The root mean square error (RMSE) was used, indicating absolute error of models; and intraclass correlation with a one-way ANOVA fixed effects model (ICC1) and its confidence intervals (95%) was used to indicate quality of the equivalence. RMSE was calculated using sklearn library (Scikit-learn, Pedregosa et al. 2011) and ICC1 using pingouin library (Vallat 2018) for Python. Both values were first calculated for each throw timeseries and for each parameter separately. Then, the mean and standard deviation values for each participant were calculated. For each model, the participant-specific values and group-specific values were presented. The low number of participants and throws must be taken into consideration while reviewing results. Based on the 95% confident interval of the ICC1 estimate, values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent concurrent validity, respectively (Koo & Li 2016).

9 RESULTS

The mean root mean square error (RMSE) varied between 7.7–16.6 degrees at the group level for all predicted parameters with the five sensor group models (table 4). Means of ICC1 for the same models indicated good agreement for bilateral hip flexion and for pelvic rotation (table 5). For front leg knee flexion and pelvic tilt ICC1 indicated moderate agreement (table 5). For pelvic list and rear leg knee flexion the ICC1 value was poor and varied a lot between participants (table 5). Examples of predictions by the five sensor group models' for all parameters are shown in figure 16.

TABLE 4. Mean and standard deviation of RMSE (absolute value as degrees) for each variable of the group model predicting all parameters using sensors on the pelvis, thighs and shanks. (F = front leg, R = rear leg.)

Subject	F-Hip flex	R-Hip flex	F-Knee flex	R-Knee flex	Pelvic list	Pelvic rotation	Pelvic tilt
1	8 ± 0.7	12.9 ± 0.4	10.4 ± 1	10.3 ± 2.4	6.4 ± 1.3	20.2 ± 1.2	6.5 ± 0.7
2	6.9 ± 1.9	8 ± 0.8	11 ± 3.4	22.5 ± 0.4	6.3 ± 1.2	8.7 ± 0.4	7.1 ± 1.5
3	6.9 ± 0.7	8.5 ± 1.2	8.7 ± 1.7	10.9 ± 1.6	7.1 ± 1.3	12.4 ± 2	5.8 ± 1.3
4	19.2 ± 1.5	12.2 ± 2	7 ± 2.8	18.3 ± 1.6	8.2 ± 2	19.3 ± 3.4	10.1 ± 0.8
5	16.6 ± 1.6	5.1 ± 0.5	10.9 ± 0.7	9.8 ± 2	5.3 ± 0.3	11.2 ± 1.2	8.1 ± 1
6	10.6 ± 1.5	9 ± 0.9	14.4 ± 3	19.6 ± 3.4	8.5 ± 1.6	13.2 ± 4.5	7.4 ± 0.8
7	8.6 ± 1.2	9.6 ± 2.3	10.6 ± 1.2	22.8 ± 1.7	4.6 ± 1.3	6.9 ± 1.8	5 ± 2
8	7 ± 1.3	15.1 ± 2.6	11.6 ± 2.2	17 ± 1.5	6.3 ± 0.8	12.6 ± 1.6	8.1 ± 1.4
9	14.9 ± 5.9	14.7 ± 1.3	18.3 ± 6.4	14.5 ± 6	11.8 ± 0.8	19.7 ± 6	10.5 ± 2.3
10	16.4 ± 2	9.6 ± 1.7	13.1 ± 3.9	20.5 ± 0.9	12.1 ± 1.6	16.4 ± 2.9	10.8 ± 1.2
	11.5 ± 4.3	10.5 ± 2.9	11.6 ± 2.8	16.6 ± 4.5	7.7 ± 2.3	14.1 ± 4.2	7.9 ± 1.8

For individual models based on five sensors the RMSEs were clearly lower for all predicted parameters compared to five sensor group models (table 6) with both ways; comparing certain individual model's accuracy to the values predicted with group model for same subject, and comparing total means of individual models or group models. The mean values of five sensor group model or later introduced two sensors knee model were not separately calculated for these five subjects whom individual models were implemented. The mean RMSEs at the individual models in group level varied between 2.4 and 5.4 degrees. At the individual level the mean RMSEs for all parameters varied between 1.5 and 6.8 degrees. Standard deviations were low and ICC1 showed excellent agreement for all parameters except pelvic list, where the ICC1 indicated good agreement (table 7).

TABLE 5. ICC1 values with lower confidence limits for each variable predicted by the group model using sensors on the pelvis, thighs and shanks. For each parameter ICC1 is on the left and the LCL (lower confidence limit) value is on the right. (F = front leg, R = rear leg)

Subject	F-Hip flex		R-Hip flex		F-Knee flex		R-Knee flex		Pelvic list		Pelvic rotation		Pelvic tilt	
1	0.91	0.88	0.85	0.8	0.73	0.66	0.53	0.42	0.47	0.36	0.77	0.7	0.82	0.77
2	0.91	0.88	0.95	0.94	0.59	0.49	-0.07	-0.22	0.43	0.31	0.94	0.92	0.74	0.67
3	0.94	0.92	0.93	0.91	0.83	0.79	0.31	0.18	0.28	0.15	0.92	0.9	0.87	0.83
4	0.62	0.52	0.89	0.86	0.86	0.82	0.21	0.07	-0.06	-0.19	0.87	0.82	0.66	0.56
5	0.59	0.49	0.98	0.97	0.86	0.81	0.74	0.67	0.42	0.29	0.9	0.87	0.63	0.53
6	0.86	0.82	0.89	0.86	0.7	0.62	-0.06	-0.18	-0.3	-0.42	0.94	0.92	0.84	0.79
7	0.91	0.87	0.92	0.89	0.76	0.69	-0.27	-0.4	0.48	0.37	0.97	0.96	0.89	0.85
8	0.92	0.9	0.77	0.71	0.73	0.66	-0.38	-0.5	-0.17	-0.31	0.92	0.89	0.72	0.64
9	0.52	0.44	0.76	0.69	0.37	0.25	0.21	0.09	-0.59	-0.68	0.58	0.49	0.46	0.35
10	0.71	0.63	0.91	0.89	0.74	0.67	0.02	-0.12	-0.33	-0.45	0.86	0.82	0.6	0.5
	0.79	0.73	0.88	0.85	0.72	0.65	0.12	0	0.06	-0.06	0.87	0.83	0.72	0.65

TABLE 6. Mean and standard deviation of RMSE (absolute value in degrees) for each variable predicted by the individual models using sensors on the pelvis, thighs and shanks. (F = front leg, R = rear leg.)

Subject	F-Hip flex	R-Hip flex	F-knee flex	R-Knee flex	Pelvic list	Pelvic rotation	Pelvic tilt
4	3.8 ± 1.9	3.1 ± 2.8	4.8 ± 1.3	4.2 ± 1.1	2.6 ± 1.7	3.2 ± 0.8	2.7 ± 2.3
5	2.6 ± 2.1	2.4 ± 1.1	4.8 ± 3	4 ± 0.8	1.6 ± 0.6	3.2 ± 1.2	1.5 ± 0.8
6	3.8 ± 0.3	3.7 ± 0.7	6.8 ± 1.4	4.4 ± 1.3	2.2 ± 0.3	5.6 ± 0.5	2.8 ± 0.4
7	2.6 ± 0.9	3.6 ± 3.1	3.9 ± 5.9	3.1 ± 1.9	2 ± 1	3 ± 6.6	2.7 ± 1.6
10	5.1 ± 0.8	3.9 ± 1	6.5 ± 2.5	4.4 ± 1	1.7 ± 1.1	3.5 ± 1	2.2 ± 0.9
	3.6 ± 1	3.3 ± 0.5	5.4 ± 1.1	4 ± 0.5	2 ± 0.4	3.7 ± 1	2.4 ± 0.5

TABLE 7. ICC1 values with lower confidence limits for each variable predicted by the individual models using sensors on the pelvis, thighs and shanks. For each parameter, ICC1 is on the left and the LCL value is on the right. (F = front leg, R = rear leg)

Subject	F-Hip flex		R-Hip flex		F-Knee flex		R-Knee flex		Pelvic list		Pelvic rotation		Pelvic tilt	
4	0.98	0.98	0.99	0.99	0.93	0.91	0.95	0.93	0.83	0.79	1.00	1.00	0.97	0.96
5	0.99	0.99	0.99	0.99	0.97	0.96	0.95	0.94	0.94	0.92	0.99	0.99	0.99	0.98
6	0.98	0.97	0.96	0.94	0.86	0.84	0.90	0.88	0.79	0.74	0.97	0.96	0.96	0.95
7	0.99	0.98	0.98	0.98	0.94	0.92	0.94	0.92	0.86	0.83	0.99	0.99	0.95	0.94
10	0.98	0.97	0.99	0.98	0.95	0.93	0.97	0.96	0.96	0.94	0.99	0.99	0.98	0.98
	0.98	0.98	0.98	0.98	0.93	0.91	0.94	0.93	0.88	0.84	0.99	0.99	0.97	0.96

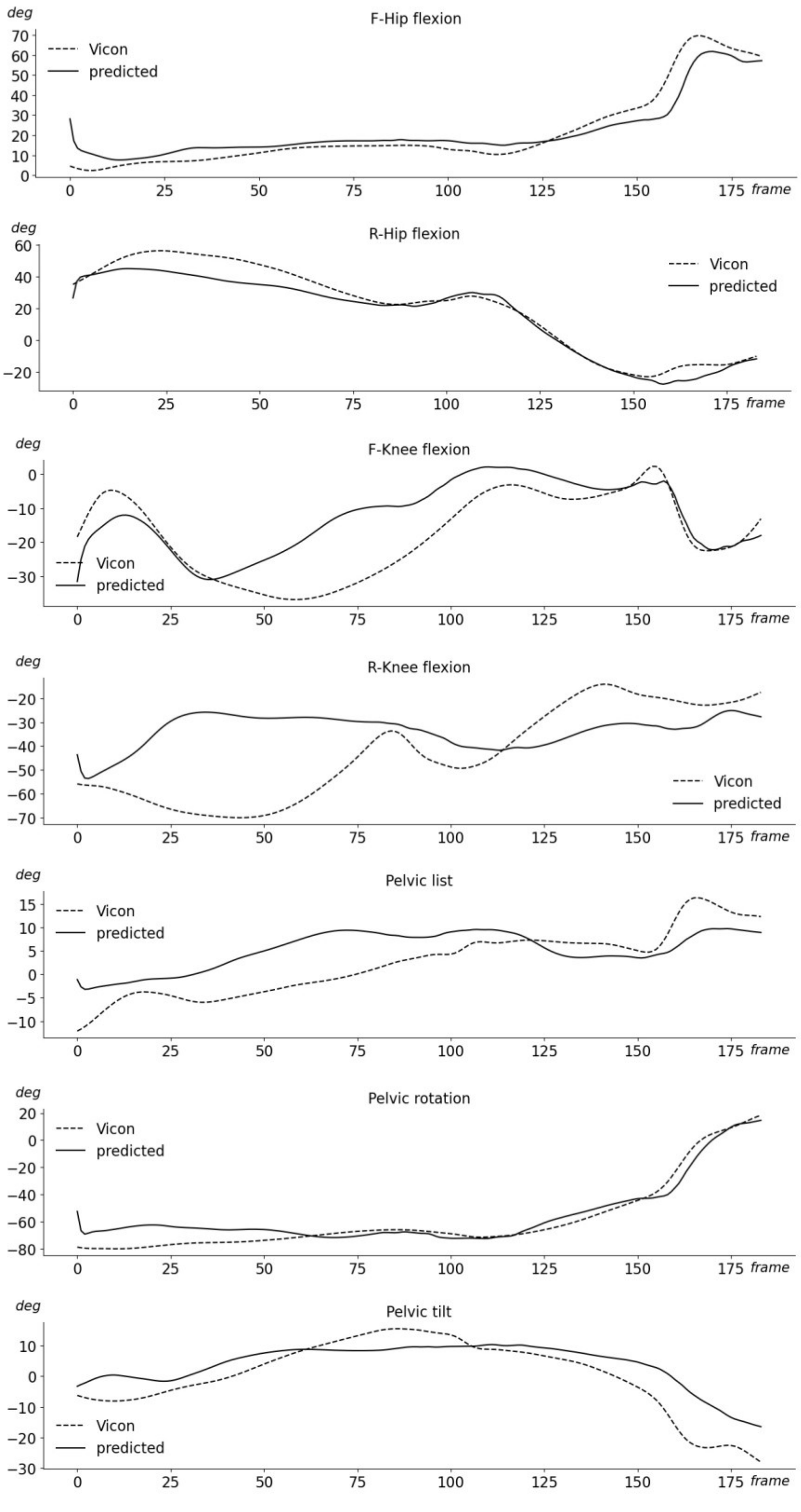

FIGURE 16. Examples of five sensor group model predictions for one participant. The curves are mean values calculated from all maximal throws for this participant. The largest absolute difference in predicted and Vicon curves was in rear leg knee flexion.

Two sensors group models for predicting front leg knee flexion showed absolute RMSEs that were between those of the five sensor group models and individual models (table 8). For subject 10, the RMSEs of predictions were on average 30.8 degrees, which is clearly more than for the same subject in the other models. Therefore, the group level RMSE was nearly the same as that of the five sensor group model. The predictions seem to be opposite in comparison to other subjects when looking at ICC1 values, so the behaviour of the model was significantly different for subject 10. ICC1s indicated excellent equivalency for three subjects, good for three, moderate for two and poor for two subjects. Figure 17 shows a comparison between three models for one participant in predicting front leg knee flexion.

TABLE 8. Mean and standard deviation of RMSE (absolute value in degrees) and ICC1 values with lower confidence limits for prediction of front leg knee flexion using thigh and shank sensors on the front leg. (LCL = Lower Confidence Limit.)

Subject	RMSE	ICC1	CI95% LCL
1	6.4 ± 0.7	0.93	0.91
2	7.6 ± 1.4	0.78	0.72
3	6.1 ± 1.6	0.91	0.88
4	6.1 ± 2.2	0.9	0.86
5	9.8 ± 2.3	0.83	0.78
6	13.8 ± 2.2	0.72	0.65
7	8.9 ± 1.1	0.81	0.76
8	11.6 ± 2.1	0.68	0.6
9	14.3 ± 5	0.51	0.4
10	30.8 ± 4.1	-0.54	-0.64
	11.5 ± 7	0.65	0.59

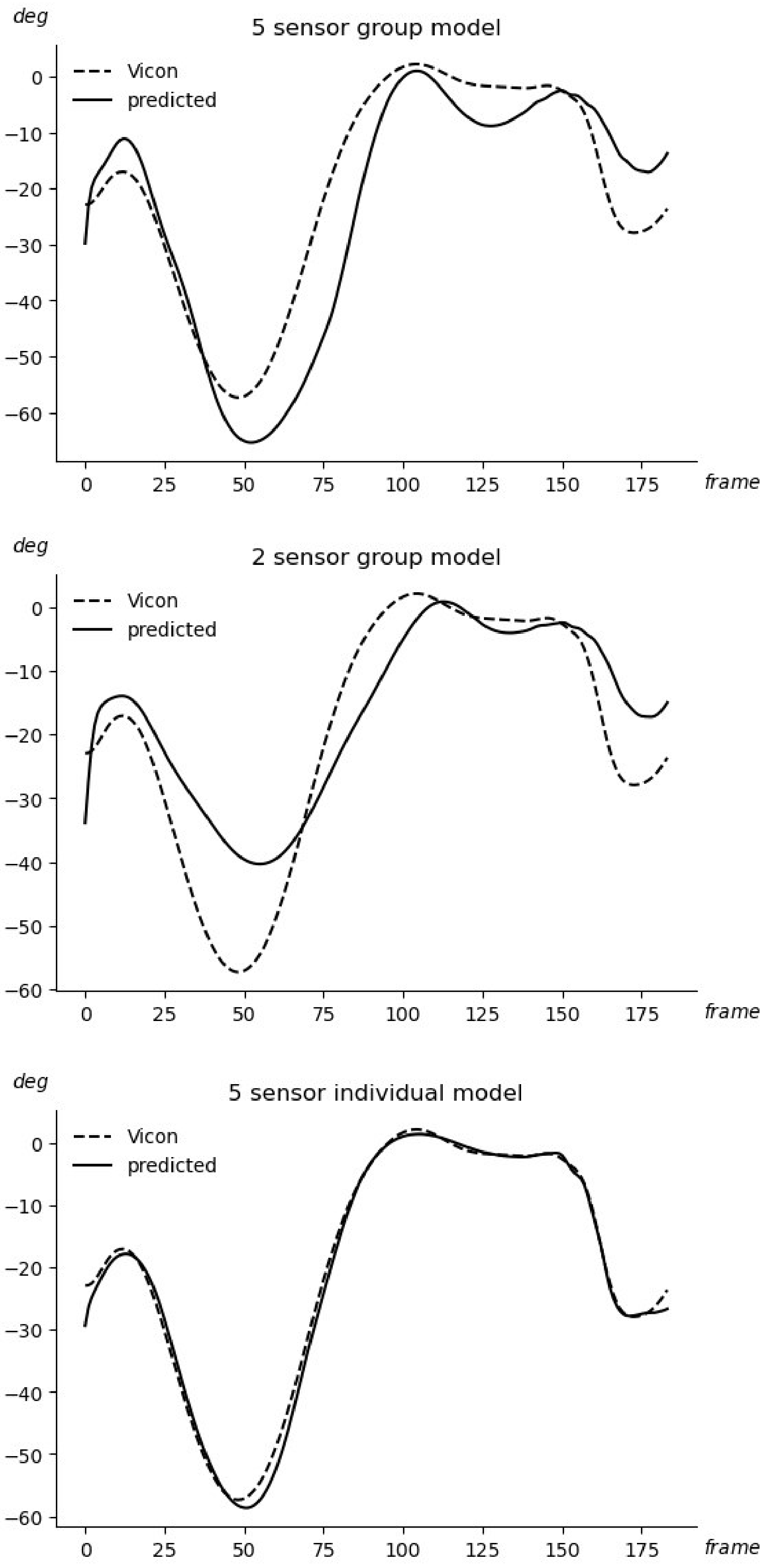

FIGURE 17. One participant's mean curves for front leg knee flexion from all models.

10 DISCUSSION

The aim of this work was to examine which kind of ML methods could be used to estimate kinematics in javelin throwing using IMU data. Long short-term memory (LSTM) models were built to predict lower body kinematics during a javelin throw performance. These models were 1) group models using five sensors and predicting all examined variables (pelvic list, pelvic tilt, pelvic rotation, rear leg hip and knee flexion, and front leg hip and knee flexion), 2) individual models using five sensors and predicting all examined variables, and 3) group models using two sensors on the thigh and shank to predict front leg knee flexion angle.

The hypothesis for the first research question was that a small sample size of only ten participants and nine to 19 throws (after data quality check) from each participant would not be sufficient to create a five sensor group model that could predict pelvic orientations and bilateral hip and knee angles with error below five degrees. This hypothesis was confirmed. The five-sensor individual models were the most accurate, and accuracy was acceptable for all variables. However, the individual models were probably overfitted, even though the submaximal throws were used in training and models were tested using only maximal throws. The second research question related to whether the final front leg knee angle could be estimated with better accuracy using a group model that only included data from thigh and shank sensors. The hypothesis was cautious because of the possibility that other sensors provide more useful information for the model. However, the two-sensor models predicted front leg knee angle better than five-sensor group models when comparing prediction accuracies separately for each athlete. The mean accuracy for the knee angle model was only slightly better and the standard deviation was larger. This was because the accuracy was clearly worse with the knee angle model than with the group model for one thrower.

Although javelin throwing is performed using the whole body and the function of the arm is important, in this study we only examined the lower body. Because most studies that have used neural networks (NN) and inertial measurement units (IMU) in combination have analysed human locomotion and kinematics during walking and running (e.g. Hernandez et al. 2021; Rapp et al. 2021), studying this area in relation to javelin throwing was easier to start with examining lower body. The predicted parameters were pelvic tilt, list and rotation, and bilateral hip and knee flexion. The period considered was from the beginning of the last crossover stride until the javelin release. Bilateral hip and knee flexion (/extension) are clearly important parameters to examine. For example, around the crucial moment of final front foot contact there is very high impact loading (Komi & Mero 1985). At this time, the functions of the front leg's hip and knee are important (Panoutsakopoulos & Kollias 2013; Whiting et al. 1991; Komi & Mero 1985). To my knowledge, the function of the pelvis has not been studied in javelin throwing separately, although Chen et al. (2020) mentioned the importance of the motion pattern starting from the front leg knee extension, which leads to hip and upper body rotation, and shoulder extension. Also, the importance of leaning the torso backwards at the beginning of the delivery phase has been discussed (Panoutsakopoulos & Kollias 2013). In some baseball studies pelvic function has been investigated (e.g. Azuma et al. 2021).

It is difficult to find boundary values for acceptable errors in a motion analysis system for each parameter because the acceptance is dependent on the purpose of use. McGinley et al. (2009) reviewed accuracies in clinical gait analysis and suggested that an error less than two degrees is highly likely to be widely considered acceptable. They described errors between two and five degrees as reasonable but requiring consideration when examining results. Errors above five degrees may lead to clinical misinterpretation. (McGinley et al. 2009.) It is probably appropriate to set acceptance levels in proportion to the range of motion. For example, pelvic tilt and list do not exhibit as large a range as pelvic rotation. When examining results, it should be remembered that Vicon and OpenSim already include errors. In this study, calculated root mean square errors (RMSEs) denote the error between Vicon's and OpenSim's "ground truth" and predictions of the built models. Thus, the real error in the predictions is higher. Built models can only be compared to "ground truth method" without knowing all of the absolute motions.

10.1 Overview of results

If we assumed an accuracy limit of five degrees RMSE, the five-sensor group model was not able to predict pelvic orientations or hip and knee flexions accurately during the timeseries. Although based on the RMSE values the error was in the same scale for all variables, the intraclass correlation coefficient (ICC1) shows that there were large differences between the variables in error variance during the timeseries. The predictions of the pelvic list and rear leg knee flexion were very poor quality. For the other variables, the model showed moderate or good agreement. Numerical observations were surprisingly good for the five sensor group model. Because there were ten different participants and training data of the model did not include data from tested participants, there was likely a reasonable amount of variation between throws used to train the model. Probably the five sensor group model started to learn the time dependencies between signals of sensors and certain angles. The angle presentations could also be presented as a change in the angle between two time points. It would be interesting to see if the estimations could be improved using this approach.

Individual models had clearly better ability to predict all parameters compared to five sensor group models. The absolute RMSE values varied between 1.5–6.8 degrees between all the variables and five participants for whom the individual models were trained. At the group level the RMSE means varied between 2–5.4 degrees for all angles. Despite the pelvic list, all the variables showed excellent agreement. For pelvic list, the agreement was good. With accuracy of individual models, the lower body kinematics could be estimated with LSTM model in javelin throwing for all variables except the final front leg knee angles, although but it is not straightforward. The models were almost certainly overfitted. Even though there can be intra-thrower variability during one throwing session (Whiting et al. 1991), probably the variation between throws was not large. It is likely that with this number of throws in the individual models (at least 14 throws in training data), the superior model accuracy was due to less variability between throws of a given thrower, compared to the variability between different throwers. Submaximal throws were used in the training data, and this makes the construction more complicated because submaximal throws likely increase the variability between throws.

At the group level, mean RMSE values produced by the model designed to predict front leg knee flexion were not clearly different to those of the five sensor group model. Nonetheless, there was a clear outlier value for one participant; knee flexion prediction with the specific model failed to predict angles for subject 10. There was a big difference compared to the other models for this participant. Thus, it is likely that there was a mistake in training the model or during other calculations. The mistake is yet to be identified, but this observation is a good example of inconstancy when examining small sample sizes. For the other participants, the error with the front leg knee flexion produced by the specific model seemed to be slightly smaller in comparison to the five sensor group models but RMSE was not under five degrees for any variable.

Lower limb sagittal joint angles (hip, knee, and ankle) can be estimated using IMUs with good accuracy (Mundt et al. 2020a). Mundt et al. (2020a) found the worst estimation accuracy in transverse plane, and frontal plane angles having a lower total range of motion had RMSEs comparable to sagittal plane. In this study only sagittal plane angles of the hips and knees were estimated. The estimation of the frontal or transversal plane hip and knee angles would likely have been less accurate. In our data, the phenomenon can already be seen when looking at prediction errors in pelvic orientation. Pelvic orientation ranges are highest in the transverse plane (pelvic rotation), second highest in the sagittal plane (pelvic tilt) and lowest in the frontal plane (pelvic list). RMSEs for these variables were largest for pelvic tilt and lowest for pelvic list, respectively, showing the RMSEs can be high in movements where the ranges are large, and predictions can have larger absolute errors (table 4 and 6). While looking at ICC1s we can see the equivalence is highest for pelvic rotation, second highest for pelvic tilt and worst for pelvic list (table 5 and 7). In the five sensor group models the predictions of pelvic list were very inaccurate. These observations support the result of Mundt et al. (2020a). The solution for this problem could be a bigger dataset including more variance in the training data.

For some throws visual observations showed that front leg knee angle estimation stabilised when achieving the rear leg contact (figure 18). If calculating results from the delivery phase the errors in predictions could probably be lower and at the same time the ICC1 agreement would increase significantly. The presented phenomenon could describe how the models built in this study learnt. The sensors had only mild variation in signal magnitudes at the beginning of the examined period because the period was cut to start when the height of the toe marker of the left foot (final front leg) started to increase. Thus, the whole take-off was not included to the examination. It may be difficult for a model to predict angles when the thrower is in the air except when the left leg (final front leg) starts to extend before the rear leg lands.

If possible, it could be advisable to use IMU signals from the whole run-up period and only predict the angles during the delivery phase. For example, the optimal knee angle at the final front foot contact leading to good kinetic energy transfer, has been proposed to be related to a controlled relationship between run up speed and an individual's skill level (Bartlett et al. 1996; Best et al. 1993; Bartlett & Best 1988). Thus, we can cautiously assume that the velocities of the segments during the run-up affect motion patterns during the delivery phase. In this thesis the option to use the whole IMU signal from the beginning of the run-up was not tested. Only simpler models were built, but building more advanced models would be the logical next step.

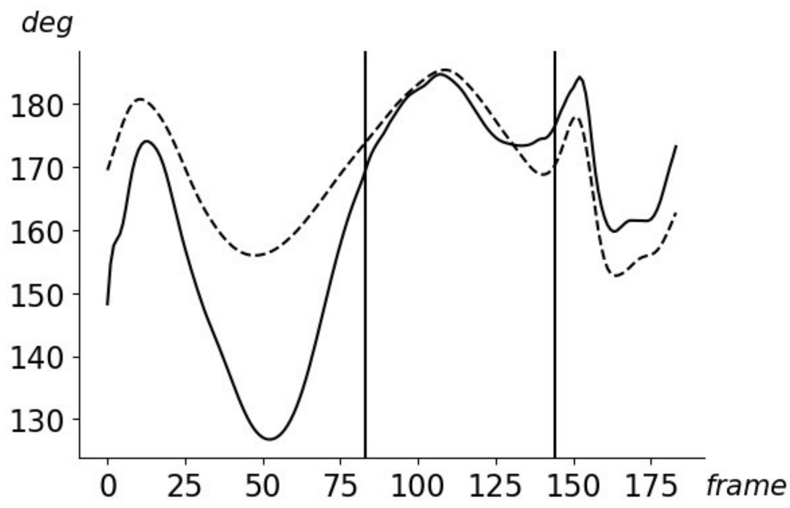

FIGURE 18. An example of front leg knee angle from one throw predicted by the five sensor group model. Angle presentation changed to illustrate absolute angle between thigh and shank segments. Vertical axes show the rear leg contact and the front leg contact, respectively. Dashed line = Vicon curve, solid line = predicted curve.

It is possible that model predictions and errors could be used to identify some performance-related issues, despite the fact that absolute predictions cannot be used to determine joint angles at a certain time point. For example, the shape of the knee flexion curve could tell how the timings of the knee angle minima are achieved during the front leg contact (Whiting et al. 1991). For this kind of use the models should be validated differently based on the errors in minimum values of knee angle in relation to the front leg contact.

10.2 Strengths and weaknesses of the study

The used sampling frequency in javelin throwing motion analysis should be high enough to detect, for example, peak segment velocities (Chen et al. 2020). In addition, a sampling frequency of 60 Hz cannot accurately detect all segment movements (Liu et al. 2014). We used a sampling frequency of 300 Hz, which is more than enough for traditional motion analysis. High sampling frequency probably helped LSTM models to learn time dependencies of motions better when there was a good number of timepoints in the timeseries. It must be remembered that the models are sensor dependent, and these models may give different results if using sensors from a different manufacturer depending on the quality of the signals.

The choice of hyperparameters was surprisingly difficult in many ways. The method of selecting the hyperparameters had to be chosen to find the most optimal hyperparameters from the separate training sets used for the tuning. Selection of the "best" hyperparameters was done by examining the means and medians of hyperparameter values for different training sets and using qualitative estimation. The full cross-validation through the data was not done because of time constraints and because structuring the data and building the leave-one-subject-out hyperparameter tuner was challenging. Surprisingly, very few examples of this type of machine learning modelling were available. Most timeseries related deep learning examples were for predicting future values, for example related to the weather or share price index. One of the weaknesses related to hyperparameter tuning was the use of a maximum of 10 epochs in the Hyberband tuner. During tuner search the number of epochs used was 50 because higher epoch values increased the running times of tuning significantly. Maybe with a higher number of epochs in tuning, the hyperparameters could be different. Later after finding the hyperparameters, the optimal number of epochs for final models was found separately.

A significant strength of our method was avoiding the use of the magnetometer data as Mundt et al. (2020c) mentioned. In an indoor environment the strength of magnetic fields can vary a lot (de Vries et al. 2009), requiring limited measurement locations or careful magnetic field mapping to increase the validity of the magnetometer signals. Also, Rapp et al. (2021) discussed that the magnetometer data often interfere with the prediction accuracy of deep learning models. It is also noteworthy that the lengths of IMU data recordings varied between throws. This could influence the amount of variety in drift in the data from different throws because the drift increases in relation to time (King 1998). However, the durations of the recordings varied from approximately 30 seconds to a couple of minutes, so the effect of drift was probably not significant.

The sacrum sensor was positioned between posterior superior iliac spines. There might be differences in how well the right place was found for each participant. For some the markers and sensor were placed on the clothing and belt. For some, the sensor was located under the clothing. This leads to variability between participants. Probably the sacrum sensor was located at the lowest lumbar vertebra. One of the aims was to predict pelvic orientations so probably more careful and similar attachment for every participant could affect the results. Still, one goal was to examine possibilities to build a model which can to some degree handle different sensor locations on a certain segment, which is also the reason why the IMU-to-segment alignment was not used.

For some athletes the reflective markers were attached on the belt and for others they were not. Also, the placements of the ankle and foot markers may have varied between throwers because those markers were attached on the surface of the shoes. Knee markers might also have introduced a soft tissue artefact and one of the throwers used a thick support bandage over the knee, which may have increased the error in marker placements. These errors in marker placement may have led to distorted training data for the LSTM models, which in turn could affect prediction accuracy. Nonetheless, these small differences in marker locations were unavoidable.

Despite the fact that prediction of frontal and transverse plane hip and joint angles is not as accurate as for sagittal plane angles (Mundt et al. 2020a), the examination of the movements in medial-lateral directions would be good in javelin throwing technique analysis (Whiting et al. 1991; Best et al. 1993). Hip adduction-abduction and hip rotation angles were possible to analyse with OpenSim using the Gait2392 model (Delp et al. 1990) but those were not included to this study. The weakness in the use of the Gait2392 model is that changes in knee angle occur only in the flexion-extension direction. Also, the highest possible knee extension value seemed to be set in the Gait2392 model based on visual observations. In javelin throwing the knee can overextend during the final front foot contact in some athletes. If more accurate motion analysis of javelin throwing with OpenSim is desirable, a different model should be used, and the system should be validated in javelin throwing against a “gold standard” method.

The indoor hall environment was a good place to measure javelin throwing performances but there were still many aspects that caused noise during data collection, especially to the Vicon system: lights, windows, other athletes in the hall and their possible reflective clothes, and furniture. One positive aspect was that the setup was placed to the hall for the whole measurement week. The cameras were not moved and that gave us the possibility to perform Vicon calibrations quickly before each test. The data was collected from seven men and three women which is not optimal for ML models because sex differences have been observed in javelin throwing (Liu et al. 2014; 2014). On the other hand, using both genders in the same models can increase the possibility to reach a generalisable model because it cannot be known if there are overlapping techniques between men and women at different athletic levels.

10.3 Suggestions for future

Our results tentatively suggest that a reduced number of sensors may be sufficient to predict joint angles. Nonetheless, there may be important information in the other sensor data, which affects for example the magnitude of the change in joint angle or the directional change in angle. Tan et al. (2022) found that the use of a single or double leg model was task dependent. It is possible to systematically test the effect of the used sensors, and to select which signals to use based on feature extraction (Sung et al. 2021) or the plane where the movement occurs and use only the signals in the plane axes (Dorschky et al. 2020). Feature selection could be done for both models: group and individual level. In addition to previous feature selection methods the different sampling frequencies, the different type of signal filtering and use of raw data could be tested. One more feature-based selection could be made testing the usefulness of adding the thrower’s weight and leg lengths to the models (Mundt et al. 2020c; Argent et al. 2019). In addition to increasing accuracy, feature selection could decrease running times in training and when using models.

When attempting to create a generalisable model, the training dataset should be representative (Dorschky et al. 2020). In this study one of the targets was to determine whether it is possible to use only data from certain thrower to predict kinematics for the same athlete. This kind of individual model cannot be generalised well if a small amount of data is used. The same athlete can develop her/his technique and subsequently the model may not recognise the angles well. Although this might suggest that a general model for all athletes would be superior, it would require a huge amount of data, which may not be feasible because of the need for "gold standard" data, which cannot be collected during competitions. Ideally, multiple countries would combine their resources and collect data using some marker-based motion capture method together with IMUs if wanting to utilise this type of IMU and ML methods.

10.4 Novelty value of this study

No previous studies have examined the prediction of joint angles in javelin throwing using IMUs and ML. This is the reason mostly walking and running studies (e.g Hernandez et al. 2021) were used when reviewing the field of kinematic estimations. Also, to my knowledge, no previous studies have predicted pelvis orientation angles with deep learning methods. This is probably because single segment orientations can theoretically be determined more easily with more general methods such as Kalman filters or Madgwick's AHRS implementation (Wada et al. 2020). The outcome of the study was promising because the variables were surprisingly well predicted with a small amount of data without a calibration pose (Zabat et al. 2019) and sensor-to-segment alignment which could reduce errors. However, the aim was to avoid using a calibration pose and sensor-to-segment alignment, which would make the method more complex, and would require accurate magnetometer data.

11 CONCLUSION

Individuality in javelin throwing technique has been highlighted in the literature (e.g. Campos et al. 2004). This suggests a need for methods to collect information about performances easily and to provide fast feedback. At the individual level athletes need feedback about ways to improve performance based on their own athletic stage (Kollias 1997, according to Manesh & Dr. Dhinu 2016). In the wider perspective, easy methods to collect more data could increase knowledge of javelin throwing at the population level and help to understand individual requirements.

Intra-thrower variability in technique, for example in front leg knee flexion-extension pattern, during one competition event (Whiting et al. 1991) suggests it could be possible to create individual models that detect the variability of a specific thrower, and could be trained to take into account changes in technique during an athlete’s development (Kollias 1997, according to Manesh & Dr. Dhinu 2016). The thrower’s development is the reason ML could be needed to build new fast methods. At the same time, the development of a certain athlete sets challenges for the models. It is not enough for the model to learn how the athlete moves. A successful model must also be able to predict new motion patterns which the model has not seen yet. This is why the group models may be superior to individual models. If training data includes athletes with enough different throwing styles, it may be possible for the model to combine patterns of different throwers to predict new subject-specific movements accurately.

The results of this study indicate that kinematics can be predicted in fast dynamic sports movements, but optimal accuracy would require long data collection with an accurate “ground truth” method. This poses several challenges and would require collaborations between different sport centres and even countries.

REFERENCES

Abadi, Martin, Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., … others. (2016). Tensorflow: A system for large-scale machine learning. In 12th Symposium on Operating Systems Design and Implementation. 265–283. Software available from tensorflow.org

Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., & Aljaaf, A. J. (2020). A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. In M. W. Berry, A. Mohamed, & B. W. Yap (Eds.), Supervised and Unsupervised Learning for Data Science (pp. 3–21). Springer International Publishing. https://doi.org/10.1007/978-3-030-22475-2_1

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. Journal of Physics: Conference Series, 1142, 012012. https://doi.org/10.1088/1742-6596/1142/1/012012

Argent, R., Drummond, S., Remus, A., O'Reilly, M., & Caulfield, B. (2019). Evaluating the use of machine learning in the assessment of joint angle using a single inertial sensor. Journal of Rehabilitation and Assistive Technologies Engineering, 6, 2055668319868544. https://doi.org/10.1177/2055668319868544

Azuma, Y., Matsui, T., Hiramoto, M., Hashimoto, R., Matsuzawa, K., Miyazaki, T., Seo, K., Watanabe, Y., Kida, N., Kai, Y., & Morihara, T. (2021). Motion Analysis Focusing on Rotational Movements of Professional Female Baseball Pitchers: Comparison with Male University Baseball Pitchers. International Journal of Environmental Research and Public Health, 18(24), 13342. https://doi.org/10.3390/ijerph182413342

Bartlett, R. M., & Best, R. J. (1988). The biomechanics of javelin throwing: A review. Journal of Sports Sciences, 6(1), 1–38. https://doi.org/10.1080/02640418808729791

Bartlett, R., Muller, E., Lindinger, S., Brunner, F., & Morriss, C. (1996). Three-Dimenslonal Evaluation of the Kinematic Release Parameters for javelin Throwers of Different Skill Levels. Journal of Applied Biomechanics, 12, 58–71.

Bergamini, E., Ligorio, G., Summa, A., Vannozzi, G., Cappozzo, A., & Sabatini, A. (2014). Estimating Orientation Using Magnetic and Inertial Sensors and Different Sensor Fusion Approaches: Accuracy Assessment in Manual and Locomotion Tasks. Sensors, 14(10), 18625–18649. https://doi.org/10.3390/s141018625

Best, R. J., Bartlett, R. M., & Morriss, C. J. (1993). A three-dimensional analysis of javelin throwing technique. Journal of Sports Sciences, 11(4), 315–328. https://doi.org/10.1080/02640419308730001

Boddy, K. J., Marsh, J. A., Caravan, A., Lindley, K. E., Scheffey, J. O., & O'Connell, M. E. (2019). Exploring wearable sensors as an alternative to marker-based motion capture in the pitching delivery. PeerJ, 1, e6365. https://doi.org/10.7717/peerj.6365

Brouwer, N. P., Yeung, T., Bobbert, M. F., & Besier, T. F. (2021). 3D trunk orientation measured using inertial measurement units during anatomical and dynamic sports motions. Scandinavian Journal of Medicine & Science in Sports, 31(2), 358–370. https://doi.org/10.1111/sms.13851

Brunetti, F., Moreno, J. C., Ruiz, A. F., Rocon, E., & Pons, J. L. (2006). A new platform based on IEEE802.15.4 wireless inertial sensors for motion caption and assessment. Conference Proceedings: EMBS Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, Suppl, 6497–6500. https://doi.org/10.1109/IEMBS.2006.260866

Caldas, R., Mundt, M., Potthast, W., Buarque de Lima Neto, F., & Markert, B. (2017). A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. Gait & Posture, 57, 204–210. https://doi.org/10.1016/j.gaitpost.2017.06.019

Camomilla, V., Bergamini, E., Fantozzi, S., & Vannozzi, G. (2018). Trends Supporting the In-Field Use of Wearable Inertial Sensors for Sport Performance Evaluation: A Systematic Review. Sensors, 18(3), 873. https://doi.org/10.3390/s18030873

Campos, J., Brizuela, G., & Ramón, V. (2004). Three-dimensional kinematic analysis of elite javelin throwers at the 1999 IAAF World Championships in Athletics. 2, 12.

Carter, J. A., Rivadulla, A. R., & Preatoni, E. (2022). A support vector machine algorithm can successfully classify running ability when trained with wearable sensor data from anatomical locations typical of consumer technology. Sports Biomechanics, 101151352, 1–18. https://doi.org/10.1080/14763141.2022.2027509

Caruso, M., Sabatini, A. M., Laidig, D., Seel, T., Knaflitz, M., Della Croce, U., & Cereatti, A. (2021). Analysis of the Accuracy of Ten Algorithms for Orientation Estimation Using Inertial and Magnetic Sensing under Optimal Conditions: One Size Does Not Fit All. Sensors, 21(7), 2543. https://doi.org/10.3390/s21072543

Chen, Y., Chou, Y.-C., Lo, T.-Y., Chang, W.-H., & Chang, J.-H. (2020). Kinematics differences between personal best and worst throws in actual javelin competition. International

Journal of Performance Analysis in Sport, 20(1), 10–18. https://doi.org/10.1080/24748668.2019.1697580

Chollet, F. (2021). Deep Learning with Python, Second Edition. New York, NY: Manning Publications.

Chollet, F., & others. (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras

Chow, D. H. K., Tremblay, L., Lam, C. Y., Yeung, A. W. Y., Cheng, W. H. W., & Tse, P. T. W. (2021). Comparison between Accelerometer and Gyroscope in Predicting Level-Ground Running Kinematics by Treadmill Running Kinematics Using a Single Wearable Sensor. Sensors (Basel, Switzerland), 21(14), 4633. https://doi.org/10.3390/s21144633

Cronin, N. J. (2021). Using deep neural networks for kinematic analysis: Challenges and opportunities. Journal of Biomechanics, 123, 110460. https://doi.org/10.1016/j.jbiomech.2021.110460

Delp, S. L.. Loan, J. P., Hoy, M. G., Zajac, F. E., Topp, E. L. & Rosen, J. M. (1990). An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. IEEE Tnsactions on Biomedical Engineering, 37(8), 757-767.

de Vries, W. H. K., Veeger, H. E. J., Baten, C. T. M., & van der Helm, F. C. T. (2009). Magnetic distortion in motion labs, implications for validating inertial magnetic sensors. Gait & Posture, 29(4), 535–541. https://doi.org/10.1016/j.gaitpost.2008.12.004

Dorschky, E., Nitschke, M., Martindale, C. F., van den Bogert, A. J., Koelewijn, A. D., & Eskofier, B. M. (2020). CNN-Based Estimation of Sagittal Plane Walking and Running Biomechanics From Measured and Simulated Inertial Sensor Data. Frontiers in Bioengineering and Biotechnology, 8, 604. https://doi.org/10.3389/fbioe.2020.00604

Elumalai, G. & Ramakrishnan, R. (2021). Development of novel and efficient approach for analyzing and monitoring the movement parameters for javelin athletes based on internet of things. Journal of Ambient Intelligence and Humanized Computing, 12, 4663–4676.

Favre, J., Jolles, B. M., Aissaoui, R., & Aminian, K. (2008). Ambulatory measurement of 3D knee joint angle. Journal of Biomechanics, 41(5), 1029–1035. https://doi.org/10.1016/j.jbiomech.2007.12.003

Fusca, M., Negrini, F., Perego, P., Magoni, L., Molteni, F., & Andreoni, G. (2018). Validation of a Wearable IMU System for Gait Analysis: Protocol and Application to a New System. Applied Sciences, 8(7), 1167. https://doi.org/10.3390/app8071167

Galbusera, F., Casaroli, G., & Bassani, T. (2019). Artificial intelligence and machine learning in spine research. JOR SPINE, 2(1), e1044. https://doi.org/10.1002/jsp2.1044

Gasparutto, X., van der Graaf, E., van der Helm, F., C., T. & Veeger, D., H., E., J. (2021). Influence of biomechanical models on joint kinematics and kinetics in baseball pitching. Sports Biomechanics 20(1), 96-108.

Geiger, R. S., Cope, D., Ip, J., Lotosh, M., Shah, A., Weng, J., & Tang, R. (2021). “Garbage in, garbage out” revisited: What do machine learning application papers report about human-labeled training data? Quantitative Science Studies, 2(3), 795–827. https://doi.org/10.1162/qss_a_00144

Godwin, A., Agnew, M., & Stevenson, J. (2009). Accuracy of Inertial Motion Sensors in Static, Quasistatic, and Complex Dynamic Motion. Journal of Biomechanical Engineering, 131(11), 114501. https://doi.org/10.1115/1.4000109

Gregor, R. J. & Pink, M. (1985). Biomechanical Analysis of a World Record Javelin Throw: A Case Study. International Journal of Sport Biomechanics 1, 73-77.

Gurchiek, R. D., Cheney, N., & McGinnis, R. S. (2019). Estimating Biomechanical Time-Series with Wearable Sensors: A Systematic Review of Machine Learning Techniques. Sensors, 19(23), 5227. https://doi.org/10.3390/s19235227

Halilaj, E., Shin, S., Rapp, E., & Xiang, D. (2021). American society of biomechanics early career achievement award 2020: Toward portable and modular biomechanics labs: How video and IMU fusion will change gait analysis. Journal of Biomechanics, 129(0157375, hjf), 110650. https://doi.org/10.1016/j.jbiomech.2021.110650

Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., & Delp, S. L. (2018). Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. Journal of Biomechanics, 81, 1–11. https://doi.org/10.1016/j.jbiomech.2018.09.009

Hernandez, V., Dadkhah, D., Babakeshizadeh, V., & Kulić, D. (2021). Lower body kinematics estimation from wearable sensors for walking and running: A deep learning approach. Gait & Posture, 83, 185–193. https://doi.org/10.1016/j.gaitpost.2020.10.026

Herring, R. M., & Chapman, A. E. (1992). Effects of changes in segmental values and timing of both torque and torque reversal in simulated throws. Journal of Biomechanics, 25(10), 1173–1184. https://doi.org/10.1016/0021-9290(92)90073-A

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Janes, W. E., Brown, J. M., Essenberg, J. M. & Engsberg, J. R. (2012). Development of a method for analyzing three-dimensional scapula kinematics. Hand 7, 400-406.

Josyula, R. 2021. A Review on Human Pose Estimation. Northeastern University, Electrical and Computer Engineering Department Augmented Cognition Laboratory (ACLab). Independent Study. Referred 18.8.2022. arXiv:2110.06877v1.

King, A. D. (1998). Inertial Navigation – Forty Years of Evolution. GEC REVIEW, 13(3), 140–149.

Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. https://doi.org/10.48550/ARXIV.1412.6980

Komi, P. V., & Gollhofer, A. (1997). Stretch Reflexes Can Have an Important Role in Force Enhancement during SSC Exercise. Journal of Applied Biomechanics, 13(4), 451–460. https://doi.org/10.1123/jab.13.4.451

Komi, P. V., & Mero, A. (1985). Biomechanical Analysis of Olympic Javelin Throwers. International Journal of Sport Biomechanics, 1(2), 139–150. https://doi.org/10.1123/ijsb.1.2.139

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine, 15(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. Informatica, 31, 249–268.

Krogh, A. (2008). What are artificial neural networks? Nature Biotechnology, 26(2), 195–197. https://doi.org/10.1038/nbt1386

Lapinski, M., Brum Medeiros, C., Moxley Scarborough, D., Berkson, E., Gill, T. J., Kepple, T., & Paradiso, J. A. (2019). A Wide-Range, Wireless Wearable Inertial Motion Sensing System for Capturing Fast Athletic Biomechanics in Overhead Pitching. Sensors, 19(17). https://doi.org/10.3390/s19173637

Lebel, K., Boissy, P., Nguyen, H., & Duval, C. (2017). Inertial measurement systems for segments and joints kinematics assessment: Towards an understanding of the variations in sensors accuracy. BioMedical Engineering OnLine, 16(1), 56. https://doi.org/10.1186/s12938-017-0347-6

Liang, F.-Y., Gao, F., & Liao, W.-H. (2021). Synergy-based knee angle estimation using kinematics of thigh. Gait & Posture, 89, 25–30. https://doi.org/10.1016/j.gaitpost.2021.06.015

Lim, H., Kim, B., & Park, S. (2019). Prediction of Lower Limb Kinetics and Kinematics during Walking by a Single IMU on the Lower Back Using Machine Learning. Sensors (Basel, Switzerland), 20(1), E130. https://doi.org/10.3390/s20010130

Liu, H., Leigh, S., & Yu, B. (2010). Sequences of upper and lower extremity motions in javelin throwing. Journal of Sports Sciences, 28(13), 1459–1467. https://doi.org/10.1080/02640414.2010.514004

Liu, H., Leigh, S., & Yu, B. (2014). Comparison of sequence of trunk and arm motions between short and long official distance groups in javelin throwing. Sports Biomechanics, 13(1), 17–32. https://doi.org/10.1080/14763141.2013.865138

Ljubisa, L., Zivota., S. & Rade, S. (2019). The Influence of Proper Release on the Javelin Flight Distance. Activities in Physical Education and Sport, 9(1–2), 23–25.

Madgwick, S. O. H., Harrison, A. J. L., & Vaidyanathan, R. (2011). Estimation of IMU and MARG orientation using a gradient descent algorithm. 2011 IEEE International Conference on Rehabilitation Robotics, 1–7. https://doi.org/10.1109/ICORR.2011.5975346

Mahmud, E. (2009). Movement Analysis for Javelin Throwers in the Qatar 2009 Championships. ISBS – Conference Proceeding Archive.

Manesh, K. M. M. & Dr. Dhinu, M. R. (2016). Kinematic parameters assessment on the javelin release performed by under twenty Indian athletes. Kinesiologia Slovenica 22(1), 1318-2269.

Mazzà, C., Donati, M., McCamley, J., Picerno, P., & Cappozzo, A. (2012). An optimized Kalman filter for the estimate of trunk orientation from inertial sensors data during treadmill walking. Gait & Posture, 35(1), 138–142. https://doi.org/10.1016/j.gaitpost.2011.08.024

McGinley, J. L., Baker, R., Wolfe, R., & Morris, M. E. (2009). The reliability of three-dimensional kinematic gait measurements: A systematic review. Gait & Posture, 29(3), 360–369. https://doi.org/10.1016/j.gaitpost.2008.09.003

Mero, A., Komi, P. V., Korjus, T., Navarro, E., & Gregor, R. J. (1994). Body Segment Contributions to Javelin Throwing during Final Thrust Phases. Journal of Applied Biomechanics, 10(2), 166–177. https://doi.org/10.1123/jab.10.2.166

Morriss, C., & Bartlett, R. (1996). Biomechanical Factors Critical for Performance in the Men´s Javelin Throw: Sports Medicine, 21(6), 438–446. https://doi.org/10.2165/00007256-199621060-00005

Mundt, M., Johnson, W. R., Potthast, W., Markert, B., Mian, A., & Alderson, J. (2021). A Comparison of Three Neural Network Approaches for Estimating Joint Angles and Moments from Inertial Measurement Units. Sensors (Basel, Switzerland), 21(13), 4535. https://doi.org/10.3390/s21134535

Mundt, M., Thomsen, W., Witter, T., Koeppe, A., David, S., Bamer, F., Potthast, W., & Markert, B. (2020a). Prediction of lower limb joint angles and moments during gait using artificial neural networks. Medical & Biological Engineering & Computing, 58(1), 211–225. https://doi.org/10.1007/s11517-019-02061-3

Mundt, M., Koeppe, A., David, S., Witter, T., Bamer, F., Potthast, W., & Markert, B. (2020b). Estimation of Gait Mechanics Based on Simulated and Measured IMU Data Using an Artificial Neural Network. Frontiers in Bioengineering and Biotechnology, 8, 41. https://doi.org/10.3389/fbioe.2020.00041

Mundt, M., Koeppe, A., Bamer, F., David, S., & Markert, B. (2020c). Artificial Neural Networks in Motion Analysis-Applications of Unsupervised and Heuristic Feature Selection Techniques. Sensors (Basel, Switzerland), 20(16), Article 16. https://doi.org/10.3390/s20164581

Needham, L., Evans, M., Cosker, D. P. & Colyer, S. L. (2021). Can Markerless Pose Estimation Algorithms Estimate 3D Mass Centre Positions and Velocities during Linear Sprinting Activities? Sensors 21, 2889. doi: 10.3390/s21082889

Norris, M., Anderson, R., & Kenny, I. C. (2014). Method analysis of accelerometers and gyroscopes in running gait: A systematic review. Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 228(1), 3–15. https://doi.org/10.1177/1754337113502472

Osis, S. T., Hettinga, B. A., Macdonald, S. M. & Ferber, R. (2014). 64–71. https://www.researchgate.net/publication/264119622.

Panoutsakopoulos, V. & Kollias, A. I. (2013). Kinematics of the delivery phase and release parameters of top female javelin throwers. Kinesiologia Slovenica 19(1), 32-43.

Panoutsakopoulos, V., Vujkov, N., Kotzamanidou, M. C., & Vujkov, S. (2016). Technique Assessment of the Javelin Relesase Performed by Young Serbian Athletes. Physical Education and Sport 14 (2), 127–136.

Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python. JMLR 12, 2825–2830.

Pathirana, P. N., Karunarathne, M. S., Williams, G. L., Nam, P. T., & Durrant-Whyte, H. (2018). IMU Sensor Dataset for human joint pose analysis [Data set]. IEEE DataPort. https://doi.org/10.21227/3BTQ-8320

Rapp, E., Shin, S., Thomsen, W., Ferber, R., & Halilaj, E. (2021). Estimation of kinematics from inertial measurement units using a combined deep learning and optimization framework. Journal of Biomechanics, 116, 110229. https://doi.org/10.1016/j.jbiomech.2021.110229

Richardson, R. T., Rapp, E. A., Quinton, R. G., Nicholson, K. F., Knarr, B. A., Russo, S. A., Higginson, J. S. & Richards, J. G. (2017). Errors Associated with Utilizing Prescribed Scapular Kinematics to Estimate Unconstrained, Natural Upper Extremity Motion in Musculoskeletal Modeling. Journal of Applied Biomechanics 33(6), 469-473.

Robert-Lachaine, X., Mecheri, H., Larue, C., & Plamondon, A. (2017). Accuracy and repeatability of single-pose calibration of inertial measurement units for whole-body motion analysis. Gait & Posture, 54, 80–86. https://doi.org/10.1016/j.gaitpost.2017.02.029

Roggio, F., Ravalli, S., Maugeri, G., Bianco, A., Palma, A., Di Rosa, M., & Musumeci, G. (2021). Technological advancements in the analysis of human motion and posture management through digital devices. World Journal of Orthopedics, 12(7), 467–484. https://doi.org/10.5312/wjo.v12.i7.467

Särkkä, O., Nieminen, T., Suuriniemi, S., & Kettunen, L. (2016). Augmented inertial measurements for analysis of javelin throwing mechanics. Sports Engineering, 19(4), 219–227. https://doi.org/10.1007/s12283-016-0194-x

Senanayake, D., Halgamuge, S., & Ackland, D. C. (2021). Real-time conversion of inertial measurement unit data to ankle joint angles using deep neural networks. Journal of Biomechanics, 125, 110552. https://doi.org/10.1016/j.jbiomech.2021.110552

Sharifi Renani, M., Eustace, A. M., Myers, C. A., & Clary, C. W. (2021). The Use of Synthetic IMU Signals in the Training of Deep Learning Models Significantly Improves the Accuracy of Joint Kinematic Predictions. Sensors (Basel, Switzerland), 21(17), 5876. https://doi.org/10.3390/s21175876

Sung, J., Han, S., Park, H., Cho, H.-M., Hwang, S., Park, J. W., & Youn, I. (2021). Prediction of Lower Extremity Multi-Joint Angles during Overground Walking by Using a Single IMU with a Low Frequency Based on an LSTM Recurrent Neural Network. Sensors (Basel, Switzerland), 22(1), 53. https://doi.org/10.3390/s22010053

Tan, J.-S., Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., Kent, P., Smith, A., O’Sullivan, P., & Campbell, A. (2022). Predicting Knee Joint Kinematics from Wearable Sensor Data in People with Knee Osteoarthritis and Clinical Considerations for Future Machine Learning Models. Sensors (Basel, Switzerland), 22(2), 446. https://doi.org/10.3390/s22020446

Trasolini, N. A., Nicholson, K. F., Mylott, J., Bullock, G. S., D.Phil, Hulburt, T. C. & Waterman, B. R. (2022). Biomechanical Analysis of the Throwing Athlete and Its Impact on Return to Sport. Artroscopy, Sports Medicine, and Rehabilitation 4(1), e83-e91.

Vallat, R. (2018). Pingouin: statistics in Python. Journal of Open Source Software, 3(31), 1026, https://doi.org/10.21105/joss.01026

Viitasalo, J., Mononen, H. & Norvapalo, K. (2003). Release Parameters at the Foul Line and the Official Result in Javelin Throwing. Sport Biomechanics 2 (1), 15–34.

Wu, X., & Zhang, X. (2017). Responses to Critiques on Machine Learning of Criminality Perceptions (Addendum of arXiv:1611.04135). arXiv: Computer Vision and Pattern Recognition, arXiv: 1611.04135. http://arxiv.org/abs/1611.04135

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. Production & Manufacturing Research, 4(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517

Whiting, W. C., Gregor, R. J., & Halushka, M. (1991). Body Segment and Release Parameter Contributions to New-Rules Javelin Throwing. International Journal of Sport Biomechanics, 7(2), 111–124. https://doi.org/10.1123/ijsb.7.2.111

Xia, P., Hu, J. & Peng, Y. (2018). EMG-Based Estimation of Limb Movement Using Deep Learning with Recurrent Convolutional Neural Networks. Artificial Organs, 42(5), E67–E77.

Zabat, M., Ababou, A., Ababou, N., & Dumas, R. (2019). IMU-based sensor-to-segment multiple calibration for upper limb joint angle measurement—A proof of concept. Medical & Biological Engineering & Computing, 57(11), 2449–2460. https://doi.org/10.1007/s11517-019-02033-7

Zago, M., Kleiner, A. F. R., & Federolf, P. A. (2021). Editorial: Machine Learning Approaches to Human Movement Analysis. Frontiers in Bioengineering and Biotechnology, 8(101632513), 638793. https://doi.org/10.3389/fbioe.2020.638793