Data Science

Hello everyone, welcome to this module on Data Science. In this module, we will look at the introduction to the data science, and then we will look at the applications and uses of data science in human movement science. So, first of all, let's see what data science is. So data science is a very multidisciplinary field which combines the concepts from computer science, mathematics and statistics, and domain-specific knowledge. So if we look at when mathematics and domain specific knowledge is being utilized or overlap, that results in our traditional research. But when computer science or information technology knowledge along with domain knowledge overlap, that leads to the software development.

And the third overlap between these three broader fields like computer science and mathematics and statistics lead to machine learning. But when all three are overlapping, that results in the domain of data science. So data science is nothing but a very multidisciplinary field which combines the knowledge from statistics, computer science, and domain-specific knowledge. And this knowledge is being then utilized to extract the valuable insights and knowledge from the data.

Various tools and algorithms are being designed using these principles to analyze and interpret very large and complex datasets. Why we need to analyze or apply these algorithms and techniques to have an informed decision-making process or to solve complex problems which otherwise are very time consuming and sometimes are prone to errors also. So in order to understand data science, let's look at the basics of data science life as we discussed previously also, but in this we will look into more detail. Data science as a field can be subdivided into four main pillars. So these pillars include domain knowledge, mathematics, and statistical skills, knowledge of computer science and communication and visualization.

Since we are talking about very large and complex datasets, this fourth pillar also plays an important role while we are dealing with the problems in data science. So first of all in domain knowledge, we must understand since we are dealing with human movement analysis or we have biomechanical data. Since human movement analysis involves biomechanical data, so we must have a understanding and knowledge of what are the different types of biomechanical data available. Second thing which we must know about is how those data points are, what are the methods which are being utilized to collect these data in biomechanics. The next domain knowledge which we need for data science is the knowledge of biomechanical model, how the different components of a biomechanical principles are being applied like for example, in a biomechanical model, we can think of how skeletal system is producing a movement with a force input from the muscular system or how different segments are related to each other to find the joint angles.

So all this information is contained in your biomechanical model. And finally, we must understand the biomechanical parameters like once we have the data, we know the source of data, we know how that data is being utilized in a biomechanical model and that will result in biomechanical parameters. So these biomechanical parameters as we discussed in our previous modules can be joint angles, joint forces or spatio-temporal parameters also. In mathematics and statistics to work in the field or to apply data science in human movement science from mathematics and statistics, we need the knowledge of linear algebra, calculus, descriptive statistics, for example, mean, mode and medium, inferential statistics like regression equations. So these will help us to analyze the or to perform mathematical computations on the data which we have.

And then we need some knowledge about the computer science. Since here we are talking a very large amount or very huge data sets. So we need to understand how those big technology, big data technologies work and then store this data. Although programming, you do not need to be an expert in programming to become a data analyst or data scientist. However, the knowledge of programming will help to implement the mathematical skills to get the required outputs.

And finally, since all these data which like, you know, we are talking about very huge data sets here. So all these data is stored in databases. So we need to have a basic understanding how that database store and how you can communicate and extract and re-save data in the databases. And finally, we need to have storytelling skills. Why we need to have storytelling skills in this because otherwise when you are going to present this data to a diverse range of audience, some of them might not be very well versed with the concepts of different domains.

So in this case, you need to connect how the knowledge of one domain can be applied into the other domain. You also need to have a better understanding of visual art design so that you can because if you just present the numbers, sometimes it gets confusing. But if you present a picture that tells more information than just the numbers and then explaining all this data to a diverse audience. For example, in case of sports science or let's apply, like see how these four pillars are important in sports science perspective. So from domain knowledge, you need to understand the various movements which you are recording and then looking at different parameters like joint angles or positions, velocities, accelerations.

Then you need to use the mathematical skills. For example, like from biomechanical data, you might just have the point data and then use the principle of linear algebra and calculus to calculate the velocities and accelerations. And finally, all this data is being stored in a database, maybe all this data in a specific organization or organization with sports different athletes, this data is stored in their athlete management system and how to communicate with that database to extract all this information. And finally, how to present this output from using this output, and finally, how we can use this output and present it to diverse

population. For example, how the information which we obtain from all these processes to an athlete or to a coach or discuss with the sports scientists for research purposes also.

So these understanding of these four pillars is very important. So now since we are talking about biomechanical data, so let us look at what are the different type of biomechanical data. So biomechanical data can be broadly classified into qualitative and quantitative data, whereas qualitative data can be further divided into nominal data and ordinal data, and the qualitative one can be divided into continuous data and discrete data. Now let us look at what is qualitative data. So qualitative data as the name indicates cannot be measured or counted in numbers.

The examples include the images or text data. So this qualitative data is further subdivided into nominal data and ordinal data. So nominal data is the data which includes variables and do not have any order, whereas the ordinal data have sequential order by their position. So nominal data sometimes also referred to as categorical data, like this means the data can be defined into subcategories, like they can be divided into subcategory 1 or 2, we will look at the example to understand it in a better way. And ordinal data is sometimes referred as in between data qualitative and quantitative.

So examples for nominal data include like gender, so the two subcategories might be male or female. Similarly, nationality where you can identify individuals under different categories according to their nationality. In case of ordinal data, so why we call it as in between qualitative and quantitative? For example, you provide a feedback on a scale of 0 to 10. Here we are recording the output in terms of numerical data like 0, 1, 2, so on to 10. However, we cannot perform any mathematical functions on it or any arithmetic functions on it.

So another example for ordinal data is the education level. So education level is also have a sequential order, like you start with your high school, secondary school, go to college, and then go to postgraduate degrees. Next, we will look at the quantitative data. So quantitative data can be measured or counted in numbers. For example, time series data.

Continuous data, sorry qualitative data, qualitative data can be further subdivided into continuous data and discrete data. So continuous data can be expressed in the form of fractions and decimals, whereas discrete data is distinct and separate values, and it can be represented only in whole numbers or integers. It can be measured, the continuous data can be measured, like we can measure the kinematic data in a biomechanical analysis, whereas discrete data can only be counted and cannot be measured. Continuous data can be divided into smaller parts, whereas discrete data cannot be divided into smaller parts. The example of continuous data include kinematic parameters such as joint angles.

For example, we studied about like how does a knee joint angle will look like from 0 percent to 100 percent, and it can be further subdivided into smaller parts. Whereas in

discrete data, the example would be like total number of players in a competition, so which cannot be divided into smaller parts. We cannot say there are like five and a half players, either it would be five players or six players. So this will give us a better understanding how quantitative data, this give us better understanding how quantitative data can be further subdivided into continuous and discrete data sets. So now, let us look at the various applications of data science in human movement science.

So we already know human movement science has wide variety of applications in our life. And now we can see with the knowledge of or with the advancement in the field of data science, how these applications can be further improved or we can incorporate data science in these applications. The first and foremost application of human movement analysis or human movement science is gait analysis, where we look at in gait analysis what we do from human movement science perspective is look at how a person walks, which is very crucial in certain sub specialties or very specific specialties such as orthopedics, where this data can be utilized for pre surgical planning and rehabilitation where the information from gait analysis can be utilized to look at the progression of healing and then how fast or what would be the loading pattern to get a better outcome. And finally in sports science for performance enhancement the information from gait analysis will help sports scientists and coaches to optimize the training program for the athletes for better performance. So now let us look how data science can be incorporated in gait analysis.

So with the tools and techniques from data science applied in gait analysis will help recognize and classify different gait patterns like if we have data from two athletes and then there are some abnormality in the movement pattern. So data science algorithms can help us to recognize and classify those different gait patterns. It also help in diagnostics and treatment planning. So in diagnostics also particularly from orthopedic point of view, so you can use data science to classify the movement pattern for different pathological disease and then that will help for treatment planning of the patients. Another example we can see of data science in human movement science is data collection and sensors.

So in human movement science we use various sensor and technologies like motion capture system, force platforms, variable devices or electromyographic sensors and these are used to collect data on human subjects. How data science can incorporate? We can incorporate data science in the collection, processing, and analysis of large data sets. So we can automate collection procedures and then synchronize them to have a better meaningful information from the raw data sensors, raw sensor data. So another example of human movement science is fall prevention and detection in elderly population. So there has been an like in fall prevention and detection.

So in fall prevention and detection, human movement science help us to understand the various biomechanical factors which are contributing towards the fall, specifically in older

population and data science can be utilized to recognize those indicative pattern of fall and enable the development of various interventions to prevent such incidents. In health monitoring and wearables, human movement science continuously look at the movement patterns or sleep quality and other physical activity levels. Whereas data science can help us to process and analyze this information which will further provide us with valuable insights of a person's overall health and wellness. In terms of neurological movement or disorders, human movement science investigate various movement disorders and look at various biomechanical parameters like the pattern of movement, how that movement is being generated or looking at the mechanics of that movement. For example, in case of Parkinson disease or stroke patients.

So, by incorporating data science and using advanced analytics, we can identify the biomarkers to identify these disease at an early stage. Since we are using lot of wearable devices in our day-to-day life like smart watches or smart phones or other smart devices, so they can measure some of these biomarkers and then help us to early identification of these disorders. These algorithms can also help in the treatment and monitoring of the rehabilitation process for these neurological disorders. In human movement science, we perform biomechanical analysis. So in biomechanical analysis, what we do is we use the principle of mechanics and apply them on the human body during a movement.

And this analysis will help us to understand various factors like the forces and torques involved and the various joint angles like which different segments are making with each other while performing that activity. So data science, specifically machine learning and statistical methods can be applied on biomechanical data to uncover patterns and correlations which will help us to understand the various variables who are interacting with each other to produce that movement. Otherwise it will be very complex task, although researchers has been doing it manually also, but with the incorporation of data science that will improve the efficiency and also result in less errors. With this I would like to thank you for your time and patience. Thank you.