

Learning data science methods through a mobile device and full body motion data

Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2023, Vol. 67(1) 277–282
Copyright © 2023 Human Factors and Ergonomics Society
DOI: 10.1177/21695067231200871
journals.sagepub.com/home/pro



SeHee Jung¹, Hanwen Wang¹, Bingyi Su¹, Lu Lu¹,
Liwei Qing¹, and Xu Xu¹

Abstract

This study presents a mobile app that facilitates undergraduate students to learn data science through their own full body motions. Leveraging the built-in camera of a mobile device, the proposed app captures the user and feeds their images into an open-source computer-vision algorithm that localizes the key joint points of human body. As students can participate in the entire data collection process, the obtained motion data is context-rich and personally relevant to them. The app utilizes the collected motion data to explain various concepts and methods in data science under the context of human movements. The app also visualizes the geometric interpretation of data through various visual aids, such as interactive graphs and figures. In this study, we use principal component analysis, a commonly used dimensionality reduction method, as an example to demonstrate the proposed learning framework. Strategies to encompass other learning modules are also discussed for further improvement.

Keywords

learning app, computer vision, undergraduate education, context-rich data

Introduction

In the past decades, data science has drawn considerable attention as a solution to uncover hidden information that leads to new insights and well-informed decision-making from big data (Wani & Jabin, 2018). To this end, data science utilizes various analytical methods and algorithms that are derived from multiple disciplines, such as statistics, computer science, and knowledge in particular application domains (Baumer, 2015). Many companies and organizations are trying to exploit data science for business opportunities and/or public good. The growing popularity of data science has created a high demand for data scientists.

Universities have responded to this demand by establishing schools, degree programs, and courses for data science to strengthen students' competitiveness in the job market (Fayyad & Hamutcu, 2021). Yet, while many undergraduate students recognize the importance of data science, many of them are reluctant to take the pertinent courses because they find difficulties in statistics, mathematics, and/or programming (He et al., 2019; White, 2019). Instructors also have their own challenges. As the course enrollment size continuously inflates, instructors' typical strategy for scaling up classes is to fill up the courses with decontextualized practices (Donoghue et al., 2021). This prevents students from having the necessary contextual knowledge to interpret data, leading them to see data merely as meaningless numbers (Davies & Sheldon, 2021).

To enhance learning outcome in data science education, a number of studies have suggested using mobile devices, such as smartphones and tablets, as supplementary learning tools (Díaz-Sainz et al., 2021; Ong et al., 2021). Nowadays, as mobile devices are readily available to students, mobile learning is considered a practical approach to improve students' learning experience inside and outside the classroom (Aljawarneh, 2020; Norman et al., 2011). On a closer view, various visual aids (e.g., interactive graphs and figures) available on mobile learning promote students' interest, curiosity, and thereby increase their motivation and engagement in learning (Angga et al., 2016; Nuanmeesri, 2021). Recent studies also indicate that students can develop social and teamwork skills by participating in learning activities together with mobile devices (Baecker, 2022; Go et al., 2022). Thus, developing mobile apps for data science learning holds enormous potential in helping undergraduate students, who are avid users of mobile technology, to cultivate their data science abilities.

¹North Carolina State University, Edward P. Fitts Department of Industrial and Systems Engineering, Raleigh, NC, USA

Corresponding Author:

Xu Xu, North Carolina State University, Edward P. Fitts Department of Industrial and Systems Engineering, 915 Partners Way, Raleigh, NC 27695, USA.

Email: xxu@ncsu.edu

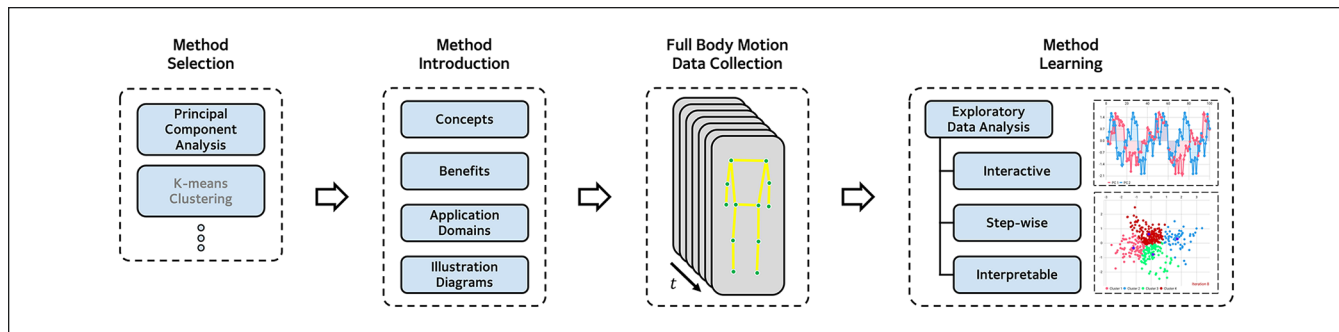


Figure 1. Four major stages of using the proposed learning platform.

Yet, there are several considerations that need to be addressed in the design of the learning app for effective data science education. First, the way concepts in data science are presented should be easy to understand in order to accommodate different levels of students. In particular, students with little knowledge and skills are more likely to lose their motivation if they find the content difficult. To reduce learning demands, the app should incorporate data visualization techniques to provide a geometric interpretation of data. This visual representation illustrates and helps students to understand how data science methods transform data into information (Unwin, 2020). Second, the dataset used for learning should be familiar and relevant to students so that they can build sufficient contextual understanding of the data. The high context-awareness of data helps students develop customized analytical approaches by enabling data-driven decision-making (Wolff et al., 2019). Thus, context-rich data could encourage students to seek different data analytical methods, which is exactly what data scientists are performing every day.

In this paper, we propose a mobile app to facilitate undergraduate students to learn concepts and methods in data science with their self-motion data. This app is specially designed to collect full body motion data via a built-in camera in a mobile device. Thus, the users can easily create datasets which are personally relevant and context-rich from their own body motion. The app also allows students to interactively explore the geometric interpretations of data with various visual aids so that they can learn data science in an effective manner.

Method

System Overview

Students will go through four stages when using the proposed app: method selection, brief method introduction, data collection through self-body motion, and method learning with the collected data (Figure 1). In the first stage, students decide which data science method they would like to study. Once a method is chosen, the app introduces the basic concepts of the method and pertinent background knowledge.

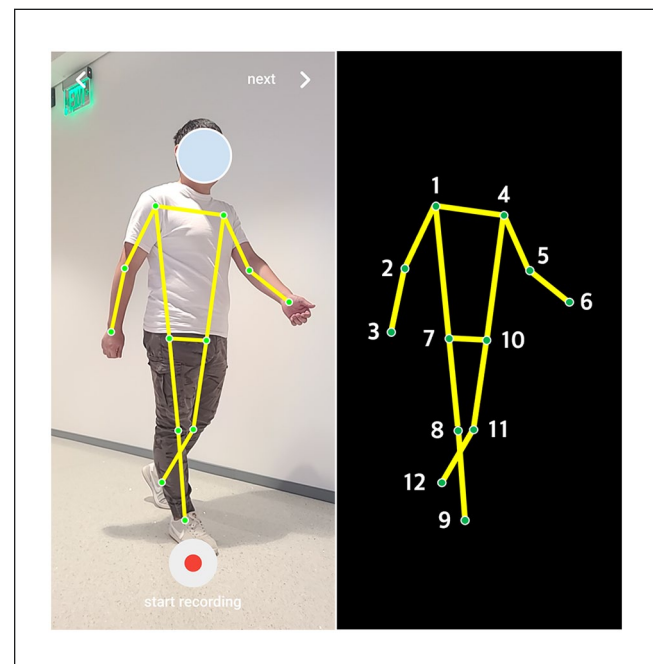


Figure 2. Learning app screen for full body motion data collection (left). Selected 12 key points of the human body (right). 1-left shoulder, 2-left elbow, 3-left wrist, 4-right shoulder, 5-right elbow, 6-right wrist, 7-left hip, 8-left knee, 9-left ankle, 10-right hip, 11-right knee, 12-right ankle.

Students are then prompted to collect their full body motion data with the help of instructional videos. During the data collection process, students can see 2D stick images superimposed on the images and compare these stick images with the human body on the screen (Figure 2). This allows students to realize that motion data are being created. Students can repeat the data collection process as many times as they want until the collected data is in accordance with what is demonstrated in instructional videos. After data collection, the app will process the data by using the selected data science method. Students can then interactively explore the geometric interpretation of the collected and processed data through a series of graphs to enhance their understanding of a data science method.

Full-body Motion Data Collection

In the past few years, the computer-vision community has developed several computation-efficient convolutional neural network (CNN) architectures for the deployment on mobile devices with limited hardware resources. For example, MobileNet and its variants (e.g., ShuffleNet; Zhang et al., 2018) use 3×3 depth-wise separable convolutions which consume approximately nine times less computational cost than standard 3×3 convolutions at only a small reduction in accuracy (Howard et al., 2017). These architectures are lightweight yet effective to extract the spatial features of the human body from local to global levels, making them suitable for position-sensitive computer vision tasks, such as pose estimation and face mesh detection.

Recently, researchers at Google presented BlazePose, a lightweight CNN architecture tailored for real-time inference of human pose on mobile devices (Bazarevsky et al., 2020). Inspired by MobileNet, BlazePose splits a standard 3×3 convolution into a 3×3 depth-wise convolution and a 1×1 point-wise convolution to perform convolution operations with much fewer parameters. In addition, BlazePose uses joint heatmaps to train a joint coordinate regression network and then discards the heatmap branch during inference (Bazarevsky et al., 2020). This selective use of heatmaps further reduces the network complexity, making BlazePose more efficient to run on mobile devices than traditional heatmap-based pose estimation models which retain joint heatmaps. From the perspective of learning experience, the real-time performance of pose estimation is essential to increase students' engagement and immersion in the data collection process (Wang et al., 2022). Thus, in this study we adopt BlazePose model due to its real-time performance, high accuracy, and high accessibility via Google ML Kit, a software development kit for deploying neural networks on mobile devices.

Once students initialize data collection, the app starts to read image frames from the embedded camera in YUV420 format. This format allows real-time display owing to its less transmission bandwidth. The app then converts the image frames into RGB bitmaps and feeds them into BlazePose at a rate of 10 frames per second. BlazePose consists of two sub-networks: a body detector performing pre-processing and a pose tracker detecting body key points. The detector standardizes the size of the inputs along with pose alignment to facilitate the subsequent key point detection. Specifically, the detector calculates four parameters for the first input frame: the bounding box for a person's face, the midpoint between the left and right hip joints, the size of the circle circumscribing the person, and the torso incline angle. Based on these parameters, the detector aligns the pose and crops the full body region from the input image. This cropped image is then reshaped to a $256 \times 256 \times 3$ square array and passed to the tracker. In the next phase, the depth-wise separable convolution layers extract the spatial features of the pre-processed image, and the heatmap-trained joint

coordinate regression network detects 33 body key points using these features. The detected key points at the previous frame are repeatedly used to align the pose of the next frame, making the subsequent pre-processing faster. Out of the 33 key points we selected 12 key points (Figure 2-right) that are necessary to represent the whole-body movements, such as walking and jumping. Each point has x and y coordinates indicating the location of a pixel in an RGB bitmap.

Learning Modules of Data Science Methods

In this section, we use principal component analysis (PCA) as an example learning module. PCA is one of the basic dimensionality reduction techniques that are commonly covered in undergraduate data science courses. The learning dataset is generated from a video in which a person walks forward. Note that this walking motion should be captured from a 45-degree view angle for 10 seconds while the person being captured is located at the center of the camera view (Figure 2-left). The collected dataset is then represented by $\mathbf{X} \in \mathbb{R}^{n \times p}$, where $n = 100$ time frames (10 seconds $\times$ 10 frames/second) and $p = 24$ features (12 key points in 2D space).

The first step of PCA is to standardize the dataset $\mathbf{X}$. Each element $\{x_{ij} : i = 1, \dots, n; j = 1, \dots, p\}$ is subtracted by its mean $\bar{x}_j$ and divided by its standard deviation s_j . We define the standardized $\mathbf{X}$ as $\tilde{\mathbf{X}}$. Then, all the features of $\tilde{\mathbf{X}}$ are mean-centered and scaled to unit standard deviation. The app then explains the benefits of standardization: in our case, we standardize the features to offset the imbalance among the joint motion variances. Because of the structural limitations of the human body, several proximal joints (e.g., shoulders, hips) have a limited range of motion compared to distal joints (e.g., wrists, ankles). As a result, the features regarding proximal joints tend to have relatively small variances, making their contributions to the principal components (PCs) less. Standardization prevents scale-variant PCA from underestimating those variables.

In PCA, we are interested in reducing the dimensionality of data. This inevitably comes at the expense of information loss, but the goal is to trade the minimum loss for the maximum simplicity by retaining most of the data structure (i.e., variance) (Shlens, 2014). From a geometric point of view, this means projecting data onto smaller dimensional subspaces that spread the data most widely. In the app, students will see a related demonstration example of data projection from 2D to 1D space, in which they will notice there is a line that maximizes the variance of the projected data. Algebraically, this can be achieved by the dot product of data and the eigenvectors of the covariance matrix Σ , which is calculated in terms of $\tilde{\mathbf{X}}$ as

$$\Sigma = \frac{1}{n-1} \tilde{\mathbf{X}}^T \tilde{\mathbf{X}} \quad (1)$$

Students need a clear understanding of the covariance matrix because it determines the eigenvectors onto which $\tilde{\mathbf{X}}$ is projected to maximize the variance and reduce the

dimensionality. Note that the concepts of a covariance matrix, an eigenvector, and an eigenvalue are described in the method introduction in advance.

Students will learn how to derive the PCs from their data $\tilde{X}$, which can be decomposed by singular value decomposition (Baker 2005) as follows.

$$\tilde{X} = USV^T \quad (2)$$

where $U \in \mathbb{R}^{n \times n}$ is an orthonormal matrix, $S = \begin{bmatrix} D \\ O \end{bmatrix} \in \mathbb{R}^{n \times p}$,

$D = \text{diag}(\sigma_1, \dots, \sigma_p)$ with singular values $\sigma_{1 \leq k \leq p} = \sqrt{(n-1)\lambda_k}$ in the decreasing order on the diagonal (i.e., $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$), $V = [v_1 v_2 \dots v_p]$, $v_k \in \mathbb{R}^{p \times 1}$ is an orthonormal eigenvector with an associated eigenvalue λ_k . Then, Eq. (1) can be rewritten by Eq. (2) as

$$\Sigma = \frac{1}{n-1} \tilde{X}^T \tilde{X} = \frac{1}{n-1} V^T [D^T \ O] U^T U \begin{bmatrix} D \\ O \end{bmatrix} V = V^T \left(\frac{D^2}{n-1} \right) V$$

and post-multiplying both sides by V , we have

$$\Sigma V = V \left(\frac{D^2}{n-1} \right) \quad (3)$$

From Eq. (3) we can see that $\Sigma v_k = \left(\frac{\sigma_k^2}{n-1} \right) v_k = \lambda_k v_k$, where v_k and λ_k are the eigenvectors and eigenvalues, respectively for the linear transformation Σ .

The app explains that the eigenvalues λ_k represent the variances of the corresponding PCs and their sum $\Lambda \equiv \sum_{k=1}^p \lambda_k$ is the same as the original variance (i.e., sum of the variances of the original features). From this fact, some students may notice λ_k / Λ represents how much proportion the k -th PC preserves the original variance. Otherwise, this will become clear through a scree plot. Students will see the scree plot where λ_k / Λ are arranged in a descending order (Figure 3) and realize the first few PCs, which are related to the first few eigenvalues, account for most of the original variance. This implies that one's full body motion during walking could be approximately explained by these two new features. The app then provides a time series of PC1 and PC2 (Figure 4). For walking motion with the instructed capturing conditions, both PC1 and PC2 form a sinusoidal-like curve that oscillates above and below zero in a periodic manner. The period of PC1 is about as half as the period of PC2. An explanation is that PC1 represents the left-right wobbling motion during walking while PC2 corresponds to the up-down motion. The up-down motion repeats every step and the left-right motion repeats every two steps. This finding is aligned with a previous study in which PCA was adopted to analyze human walking data (Daffertshofer et al., 2004).

Discussion

Despite the great potential of mobile learning, there is only limited work on developing mobile learning platforms for

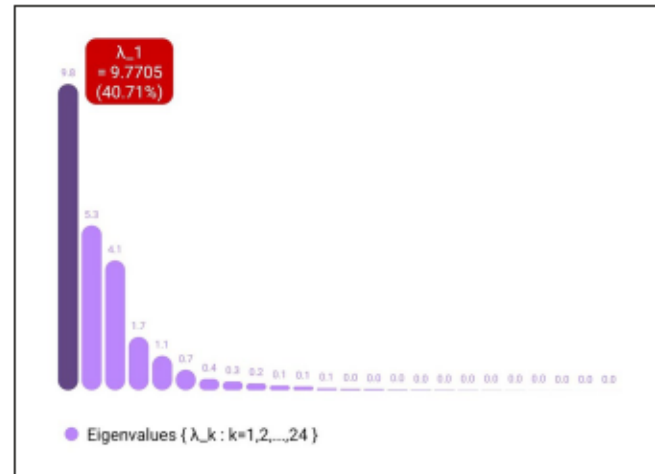


Figure 3. An example of the interactive scree plot. Students can check the value of λ_k and how much proportion the corresponding PC accounts for the original variance when touching on each bar.

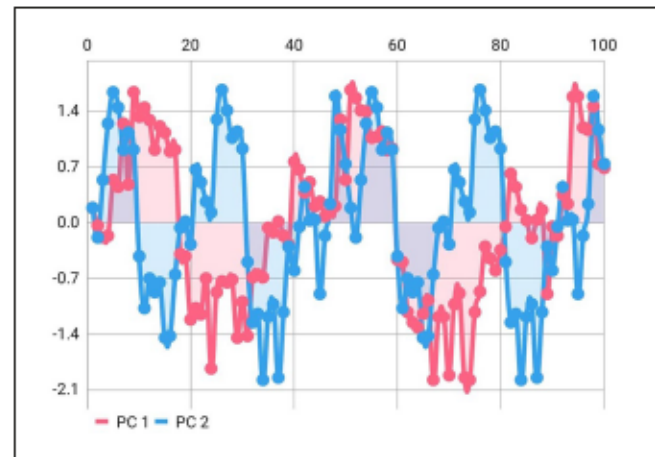


Figure 4. PC1 and PC2 over time for a walking motion.

data science. In this work, we sought to fill this gap by developing a dedicated mobile learning app for undergraduate data science education. The proposed learning app could engage and motivate students to learn data science with features that students cannot otherwise experience in traditional settings (e.g., interactive visual aids and self-data collection). Therefore, introducing this app in data science courses could complement the existing classroom curriculum. We expect that our app will serve as a supplementary learning tool for undergraduate data science courses with which students can generate personally relevant and context-rich data.

At this time, we only developed one learning module, which is for learning PCA. To cover a wide range of methods in data science, new learning modules will be introduced to the app. These modules will cover methods commonly used in data science. Following the PCA module, the next module

we could develop is a k-mean clustering module. The k-mean clustering module will guide students to pose several body postures. Among these postures, some are similar to each other (e.g., walking vs. running) while some are not (walking vs. drinking). Each posture will be captured multiple times. The dimensionality of the posture datasets will be first reduced to two (i.e., PC1 and PC2) by PCA for data visualization. Students will then see that the resulting PCs of two similar postures also have similar values and thereby form a cluster in a scatter plot. In addition, students will also notice that some less similar postures could be clustered together if the parameter k , the number of clusters, is less than the number of the performed postures.

Real-world data analysis rarely ends with a single method. Rather, it is a common practice to analyze data with a series of methods, each of which has its own role and facilitates the next step. For example, in the k-mean learning module above-mentioned, k-mean is applied on PCs. These PCs is derived from the PCA learning module. Since the first two PCs only contain partial but critical information of the original posture data, applying k-means clustering can be computationally efficient and still achieve high performance. We believe such a connection between methods is as critical as understanding how each method works. Therefore, we plan to develop interconnected learning modules so that students can cultivate their thinking ability between methods, experience the entire analysis process, and consolidate what they have learned.

In addition, since the full-body motion data are high-dimensional, dimension reduction techniques provide a great remedy for turning the data into visible 2D dataset. Therefore, we will keep exploiting dimension reduction techniques as pre-processing methods for data visualization and delivering other pertinent methods.

There are several limitations in the current app to be addressed. First, because the adopted pose estimation model is specialized in capturing a single person without occlusion, students should avoid the scenarios where there are multiple people or objects block the person of interest when they collect human motion datasets. Otherwise, the key point detection algorithm may detect the motion of another person or may be less accurate because of the view block. Second, given limited hardware resources of mobile devices, real-time inference of 3D human pose is yet difficult to implement. Compared to 3D pose analysis, the current 2D localization results could be more deviated from the actual locations due to the absence of depth information, which is the reason that the proposed app currently guides students to follow certain capturing conditions. We leave the improvement of the pose estimation model as future research along with new learning modules.

Acknowledgments

This manuscript is based upon work supported by the National Science Foundation under Grant # 2013451.

ORCID iDs

SeHee Jung  <https://orcid.org/0000-0001-7980-3911>

Bingyi Su  <https://orcid.org/0009-0008-0806-8890>

References

- Aljawarneh, S. A. (2020). Reviewing and exploring innovative ubiquitous learning tools in higher education. *Journal of computing in higher education*, 32, 57-73.
- Angga, N. M., Adelia, T., & Soetantyo, I. (2016). Enhancing Science Learning Experience By Adopting 3d Model In Mobile Learning Application—Case Study On Senses And Blood Circulatory System.
- Baecker, D. (2022). Introducing Audio Podcasts into a Practical Laboratory Course for Pharmacy Students as a Novel Tool for Performance Assessment. *Pharmacy*, 10(2), 40.
- Baker, K. (2005). Singular value decomposition tutorial. *The Ohio State University*, 24.
- Baumer, B. (2015). A data science course for undergraduates: Thinking with data. *The American Statistician*, 69(4), 334-342.
- Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). BlazePose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204*.
- Daffertshofer, A., Lamoth, C. J., Meijer, O. G., & Beek, P. J. (2004). PCA in studying coordination and variability: a tutorial. *Clinical biomechanics*, 19(4), 415-428.
- Davies, N., & Sheldon, N. (2021). Teaching statistics and data science in England's schools. *Teaching Statistics*, 43, S52-S70.
- Díaz-Sainz, G., Pérez, G., Gómez-Coma, L., Ortiz-Martínez, V. M., Domínguez-Ramos, A., Ibañez, R., & Rivero, M. J. (2021). Mobile learning in chemical engineering: An outlook based on case studies. *Education for Chemical Engineers*, 35, 132-145.
- Donoghue, T., Voytek, B., & Ellis, S. E. (2021). Teaching creative and practical data science at scale. *Journal of Statistics and Data Science Education*, 29(sup1), S27-S39.
- Fayyad, U., & Hamutcu, H. (2021). How can we train data scientists when we can't agree on who they are. *Harvard Data Science Review*, 3 (1).
- Go, M., Golbin, R., Jr, Velos, S., Dayupay, J., Dionaldo, W., Cababat, F., . . . Ocampo, L. (2022). Evaluating digital mathematical games in improving the basic mathematical skills of university students. *International Journal of Mathematical Education in Science and Technology*, 1-23.
- He, X., Madigan, C., Wellner, J., & Yu, B. (2019). Statistics at a crossroads: Who is for the challenge. In *Workshop report. National Science Foundation*. https://www.nsf.gov/mps/dms/documents/Statistics_at_a_Crossroads_Workshop_Report_2019.pdf.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., . . . Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Norman, H., Din, R., & Nordin, N. (2011). A preliminary study of an authentic ubiquitous learning environment for higher education. *learning*, 3(4), 89-94.
- Nuanmeesri, S. (2021). Developing gamification to improve mobile learning in web design course during the COVID-19

- pandemic. *International Journal of Information and Education Technology*, 11(12), 567-573.
- Ong, J. S. H., Mohan, P. R., Han, J. Y., Chew, J. Y., & Fung, F. M. (2021). Coding a Telegram Quiz Bot to Aid Learners in Environmental Chemistry.
- Shlens, J. (2014). A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*.
- Unwin, A. (2020). Why is data visualization important? what is important in data visualization?. *Harvard Data Science Review*, 2(1), 1.
- Wang, H., Xie, Z., Lu, L., Su, B., Jung, S., & Xu, X. (2022). A mobile platform-based app to assist undergraduate learning of human kinematics in biomechanics courses. *Journal of Biomechanics*, 142, 111243.
- Wani, M. A., & Jabin, S. (2018). Big data: issues, challenges, and techniques in business intelligence. In *Big Data Analytics: Proceedings of CSI 2015* (pp. 613-628). Springer Singapore.
- White, D. (2019). A project-based approach to statistics and data science. *Primus*, 29(9), 997-1038.
- Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy activities to support children's inquiries from complex data. *International Journal of Human-Computer Studies*, 129, 41-54.
- Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6848-6856).