A Statistical Size Recommender for Safety Footwear Based on 3D Foot Data

Jorge VALERO [1], Alfredo BALLESTER [1], Beatriz NACHER [1], Clara SOLVES [1], Cataldo DE LUCA [2], Marta PESCE [2], Sandra ALEMANY [1]
[1] Instituto de Biomecánica de Valencia, Valencia, Spain;
[2] Base Protection S.r.l., Barletta, Italy

https://doi.org/10.15221/23.40

Abstract

Shoe size recommendation remains a significant challenge for the footwear industry. Getting a shoe that does not fit leads to customer dissatisfaction and high return rates. In the case of safety footwear, this challenge is even greater because wearing the wrong size can compromise the safety of workers. It is thus crucial to develop a technology that allows industry to efficiently provide their employees with the right shoe size advice, in a fast, simple, and effective manner. This paper describes the results of the joint cooperation between IBV and Base Protection S.r.l. to develop and deliver such a system.
The proposed technology uses low-cost 3D scanning technology (Domescan/IBV and 3Davatar feet) to accurately capture 20 foot features and a shoes size recommender based on Multinomial Logistic Regression (MLR). The system was trained for 14 shoe lasts and used data from fitting tests of 60 subjects from both genders. It was validated with fitting tests of 25 subjects achieving an 60-80% success rate in recommendations, depending on the shoe model. The results also showed that personal fit preference plays a crucial role in size selection, hindering greater accuracy. In this regard, one of the main advantages of MLR is its informative output, i.e. a map of fitting probabilities for each size, which offers multiple options for the development of the user interface layer and may enable that the final consumer to make an informed decision based on it. The system also included an insole recommender (low, mid and high arch) that uses a classic two-dimensional recommendation grid based on foot arch indexes based on three foot features. These technologies were embodied into a physical booth for brick-and-mortar stores and into an app that directs the consumer to the nearest point of sale. This system represents a significant advance in the footwear industry and offers a streamlined solution for brands and retailers. Overall, this work demonstrates the effectiveness of utilizing MLR in a footwear recommender system, and highlights its potential for footwear brands and retailers to reduce returns and increase sales, for consumers to get a better comfort and safety at work, and for industries needing for safety shoes to reduce the burden of managing the footwear orders to its employees.

Keywords: safety footwear, shoe fitting, size recommendation, 3D foot, logistic regression, anthropometry

1. Introduction

Ensuring a proper fit of footwear is of utmost importance, as it significantly impacts both comfort and overall foot health [1,2]. Footwear sizing has long been approached through size charts based on foot length and, in some cases, also width measurements. However, such simplistic methods fail to account for the complex variations in individual foot morphology, asymmetries, personal preferences, etc. (figure 1), and also to the burden shoe size labeling; i.e. shoes labeled with the same size number from different brands (or even within the same brand) have an inner space that can vary significantly in length and shape (figure 2). For instance, the differences in foot length or ball girth between feet using the same size or between shoe lasts labeled with the same size number can be up to 25 mm, which is more than 3 FR/EU sizes or 2 full UK/US sizes.

Figure 1. Superimposed treads of 10 male feet wearing size 42 of a given shoe model.

Figure 2. Superimposed treads of 8 shoe lasts of size 38 from different shoe brands (right).

In the context of online purchases, the final consumers lack the opportunity for a physical try-on, hence there is an increased uncertainty about choosing the right size. Getting a size that does not fit contributes to a surge in return rates, causing significant logistical and financial challenges on retailers and hindering the seamless shopping experience customers expect [3]. This, in turn, leads to increased customer dissatisfaction, potentially tarnishing the brand's reputation and impacting in future sales. The exponential growth of online shopping has spurred the development of size recommendation systems with the primary aim of reducing product returns and stock, enhancing customer satisfaction and, in summary, developing a new market positioning strategy.

In the context of delegated safety shoe procurement made by health & safety departments, the stakes are significantly higher, because the problem not only impacts on short-term comfort but can also compromise the safety of workers and long-term musculoskeletal disorders (MSD) [4], [5]. A worker wearing improperly fitted safety shoes is more susceptible to slips, trips, or falls (due to reduced stability and support) which could have severe consequences in hazardous work environments. Moreover, a worker that is provided with uncomfortable shoes is more likely to not wearing them during all the shift.

In response to this pressing issue, some research efforts have emerged, leveraging diverse disciplines such as machine learning, data mining and biomechanics to tackle the challenges of footwear size selection, resulting in an array of sizing recommendation algorithms and methodologies [6]–[8].

On the other hand, the use of 3D scanning technologies has been increasingly adopted as a novel instrument for obtaining body metrics [9]. The adoption of these technologies by the footwear sector has fostered the generation of large databases of 3D scans from men and women worldwide [10]. This abundance of data provides valuable insights both for product design purposes and for the development of shoe recommendation systems [11]–[14]. This has unlocked new avenues for innovation, enhancing both the design and consumer experience in the realm of footwear. This adoption brings to brands and retailers the opportunity to establish stronger customer trust, loyalty, and a reputation and to deliver a new smooth and satisfying shopping experience.

However, size recommendation systems face significant challenges in their quest to provide accurate and personalized suggestions. Firstly, the diversity of foot morphologies and individual user preferences makes it challenging to find a one-size-fits-all approach [15]. Secondly, obtaining high-quality and comprehensive shoe fitting data (foot-footwear interaction) from a diverse global population is challenging because user feedback can be subjective and inconsistent. Gathering such feedback is also costly because a user feedback beyond the purchase requires interrupting the buying process, extending it, or designing specific testing/trials aimed at getting it. Finally, the continuous renewal and modification of shoe models imposed by fashion trends makes it necessary for the recommendation systems to be scalable and to adapt to stay up-to-date.

Shoe size recommendation is an ongoing and formidable challenge for the footwear industry. Overcoming the aforementioned obstacles is essential to ensure that size recommender delivers accurate, reliable, and satisfactory results.

This paper describes the results of the joint cooperation between Instituto de Biomecánica de Valencia (IBV) and Base Protection S.r.l. to develop and deliver a size recommendation system for safety footwear and insoles. IBV provided the foot digitizing technologies (i.e. *Domescan/IBV* and *3Davatar feet* [9], [16], [17]) and developed the algorithms for size and insole recommendation. Base Protection is a global safety footwear manufacturer and was responsible for integrating the algorithms and digitization technologies into their business model as the *SCAN&FIT* system.

2. Materials and methods

2.1. Data gathering

Shoe sample: Base Protection offers over 200 footwear models that are manufactured with 14 different shoe lasts, some of them are just for male sizes (39-48), some just for female sizes (35-42) and some are unisex and cover the whole size span (35-48). A set of different representative shoe models, one model per last, was selected to be used during the experimentation. Models corresponding to female lasts were provided in sizes 36-37-38, male in sizes 41-42-43 and unisex ones on both size sets.

Subject sample: A total sample of 60 subjects participated in the size recommendation study (36 men and 24 women). The participants were chosen satisfying the following profile: regular users of safety shoes in working age (18-65 years) with usual size of safety footwear of 37 (women) or 42 (men).

Development of the experiment*:* Each subject tried on six different shoe models in the three sizes following a balanced design that ensured that all models were tested a similar number of times (figure 3). Users tried each model in one foot, always the same, according to their preference. After each trial, users were asked about fitting and size assessment. General fit preference was also gathered as binary scale (if they preferred "perfect fit or slightly tight" or "perfect fit or slightly loose"). Users' feet were measured three times using *DomeScan/IBV* foot scanner. In this way, a dataset of more than 1000 records was gathered.

Figure 3. Shoe sample and experimental study.

2.2 Algorithm

Input variables: The algorithm employs a set of 20 foot measurements obtained using DomeScan/IBV or 3D avatar feet. The set of measurements used are listed in Table 1. Shoe model (shoe last) was treated as a categorical variable to parametrize the shape differences between them.

Table 1. Input measurements.

Abbreviation	Designation		Abbreviation	Designation
FL	Foot Length		**MBH**	Medial Ball Height
TP	Toes Position		**LBH**	Lateral Ball Height
TG	Toes Girth		**BW**	Ball Width
TH	Toes Height		**IG**	Instep Girth
TW	Toes Width		**IH**	Instep Height
P1M	Position of 1st Metatarsal head		**IW**	Instep Width
P5M	Position of 5th Metatarsal head		**NP**	Navicular Position
BP	Ball Position		**NH**	Navicular Height
BG	Ball Girth		**HW**	Heel Width
BH	Ball Height		**IHG**	Instep to Heel Girth

Data reduction: First, in order to reduce the dimensionality of the input, instead of picking a reduced number of variables based on expert criteria, we used Principal Component Analysis (PCA) to minimize the loss of information in this step. The components were extracted from IBV's own dataset of nearly 800 foot scans [18]. We reduced the 20 foot features to 9 principal components (PC).

Data projection: In order to optimize the data collection and the management of the samples, and to maximize the number fit trials per size, we gathered fit data focused on three consecutive sizes: 36-38 for female models, 41-43 for male models and both for unisex models. We thus needed to extend or generalize the data to be applied to the whole size span, e.g. 8 sizes in the case of females and 10 in the case of males. This can be achieved considering that shoe lasts are upscaled and downscaled geometrically from a reference size using grading rules. Different methods can be employed to estimate such grading for each foot feature of each observation. In this case, we used IBV's own database of foot scans [18] and the theoretical grading of foot length to estimate them (ISO/TS 19407:2015 - Footwear — Sizing).

Calculation of size fit probabilities: Classification is the process of predicting the class or category of given data points. The size recommendation can be seen as a classification problem if the shoe fitting is simplified into three discrete categories (tight, right and loose). The resulting predictive classifier should identify to which of these categories (output) belongs a new observation (input). Although there was a wide range of possible classification algorithms, in this case the well-known Logistic Regression was used. Specifically, since there were more than two possible discrete outcomes, the more general

Multinomial Logistic Regression (MLR). To select the best features and to estimate the best coefficients for each shoe model, stepwise method was used. IBV has already used logistic regression models to tackle this problem successfully in clothing [19], [20]. In this project, we also used data projection to bring all the observations to a single size by upscaling and downscaling the foot features characterizing each foot, making it work in all available sizes. This approach exploits the nature of the data collected (figure 4). Self-reported fit preference was not included in the model because it did not improve the success rate.

The advantage of the multinomial regression (for size recommendation) when compared to other classifiers resides in its informative output, which provides the probability of a size being tight, right and loose. Figure 5 shows an example of the probability curves for one participant. This map of probabilities (figure 6), offers many options to create a user interface to help a final consumer to find the right size. As a simplified example, the recommended size could be that one with highest right fit probability (red box in figure 6). Alternatively, two sizes could be displayed showing the probability for a compromise of tight fit or loose fit. The output of the model can also be used to provide a graphical representation of the shoe fit along the shoe size span using three fit regions: tight fit, optimal fit and loose fit (figure 7).

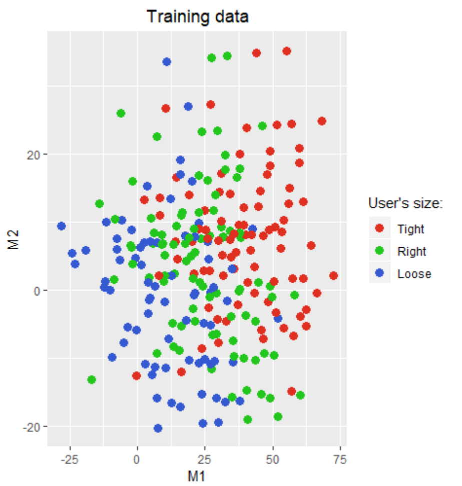

Figure 4. Training data used to fit the MLR.

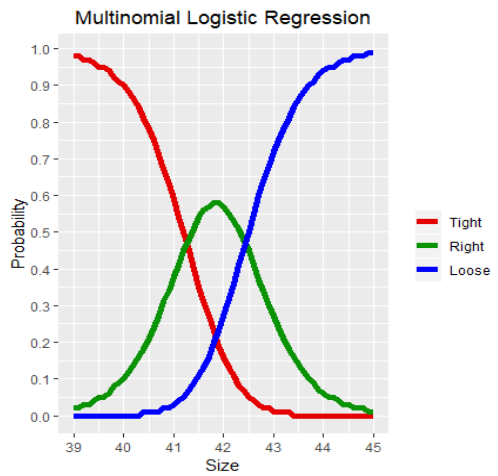

Figure 5. Probability curves for one participant.

Last Size	Tight	Right	Loose
39	1,00	0,00	0,00
40	1,00	0,00	0,00
41	0,97	0,03	0,00
42	0,87	0,13	0,00
43	0,53	0,43	0,04
44*	0,13	0,58	0,30
45	0,01	0,26	0,73
46	0,00	0,06	0,94
47	0,00	0,01	0,99
48	0,00	0,00	1,00

Figure 6. Output example.

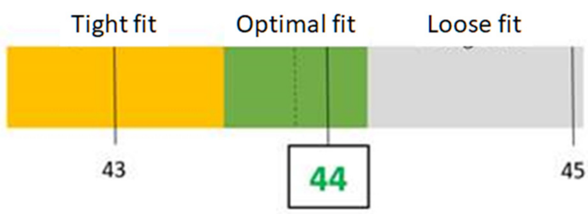

Figure 7. Example of three fit regions.

<u>Insole recommendation algorithm:</u> This algorithm was built using a database of 39 pairs of healthy feet with no pathologies, where each foot was classified as low, mid or high arch by a group of podiatrists. The recommendation algorithm was based on the foot arch capacity, which was characterized using two indices that were calculated using three foot measurements: arch area, toes position and instep height (figure 8). This way it is possible to classify feet into three categories according to three types of insoles: low arch, medium arch and high arch.

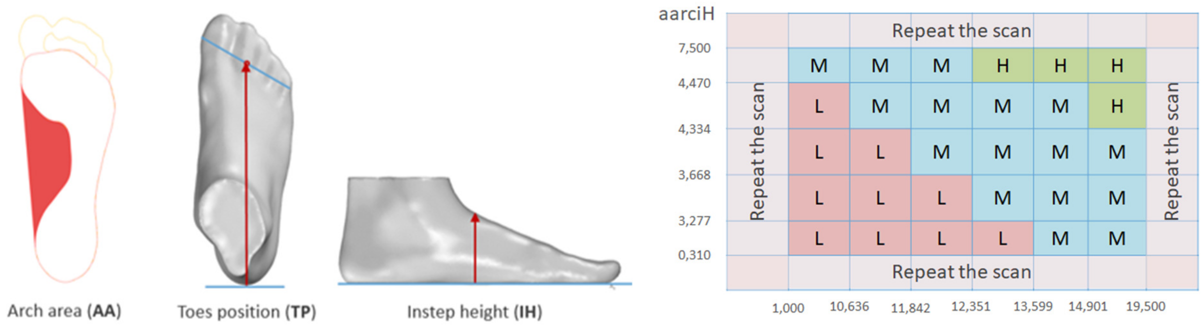

Figure 8. Measurements of the foot that are used in the arch indices' calculation.

2.3. Validation study

Shoe sample: the same 14 shoe models used in the first experiment where provided in all the size span.

Participants: A total sample of 25 users (18 men and 7 women) participated in the study. The participants satisfied the following profile: regular users of safety shoes aged between 18 and 65.

Development of the shoe size experiment: The experiment tried to simulate a real shopping experience. First, the subject's foot was acquired using DomeScan/IBV foot digitizer. A recommended size was determined using the algorithm for all shoe models. The participants tried all the models in random order for a few minutes, always starting to try each model in the recommended size. After each trial, users reported about their tolerance (capacity to wear the tested size). When trying every model, they were allowed to request and try on different sizes. Finally, users reported the preferred size for each shoe model tested.

Development of the insole height experiment: The subjects were called for a second session where they tried 3 shoe models with the recommended insole height. They knew that there were other insoles available and were allowed to request them and tried them. Finally, users reported the preferred height for each insole tested.

Key performance indicators (KPI): Fit tolerance was used as main KPI to assess results. This indicator is the percentage of users that reported that were able to wear the recommended size.

2.4. Integration of the recommender system and market release

Since base protection does not sell directly to consumers, after validation phase, the technology was embodied into a physical booth based on *Domescan/IBV* and mobile phone app for iOS and Android using *Avatar3D feet* API (figure 9, [9]). Both technologies are based on a data-driven 3D reconstruction of a foot from three calibrated images of the foot (inner, outer and top views). The former is a booth with one camera and two mirrors and the latter is an API webservice that uses as input three images taken with a smartphone. In the app case, the phone sensors and a paper sheet under the foot are used to calibrate the images. These technologies were developed by IBV and can be licensed.

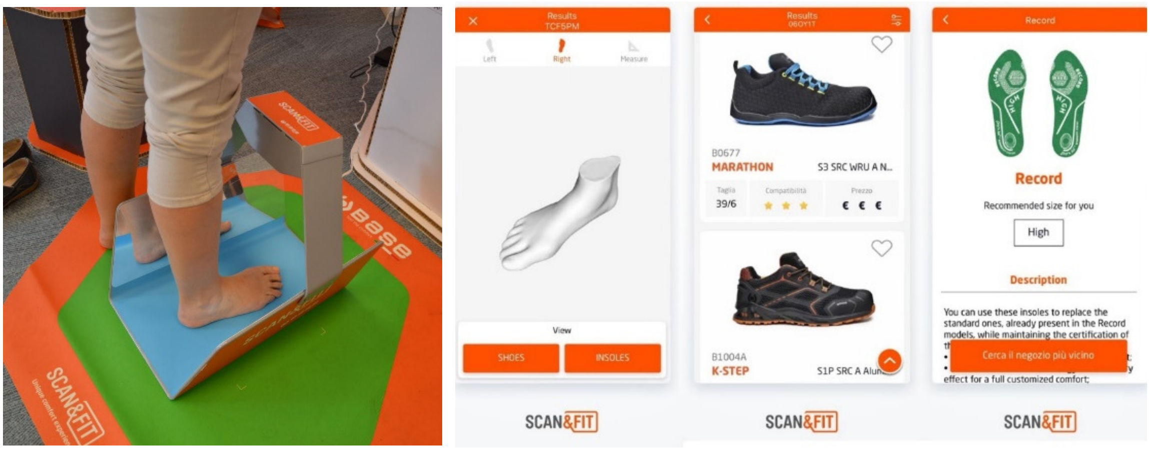

Figure 9. SCAN&FIT booth and app.

The technology was branded by Base Protection as SCAN&FIT. The interactive catalogue of safety shoe models and safety features is integrated into the technology making it easier to filter and find the suitable safety features, the optimal fit and the most appealing aesthetics among the over 200 safety shoe models that the brand offers.

In both embodiments, the process starts by digitizing the user's feet (or just one foot). Then the interactive catalogue is displayed showing for each shoe model, the suggested size for that user in each model and also a star rating of the fit provided. The rating is computed using the information provided by the logistic regression algorithm. The shoe models that are not available for the foot size of the user, e.g. female models not listed for subjects using sizes beyond size 42 and vice versa, male models are not listed for subjects using sizes below size 39.

The physical booth is aimed to be distributed in points of sale where Base Protection shoes are sold. It provides a new safety shoe buying experience making it faster to find the best Base Protection shoe. Moreover, both the booth and the experience make the brand to stand out among the other brands available at the point of sale.

A few units of the booth are embodied a portable fashion and are used by Base Protection fit technicians when they visit companies requesting for bulk safety shoe procurement. In such cases, the workers can be measured very efficiently in one visit per location/factory. The booth is also a valuable marketing asset in safety footwear international trade fairs (figure 10).

The phone app is aimed to final consumers. They can scan their feet and use it as an interactive catalogue to pick the shoe model with the right safety/technical features and know the optimal size and insole type for their feet. Then, it can be used either to refer the model and size to the health and safety department of the worker's company or to find the nearest shop where the model is available.

Figure 10: SCAN&FIT booth in a fair.

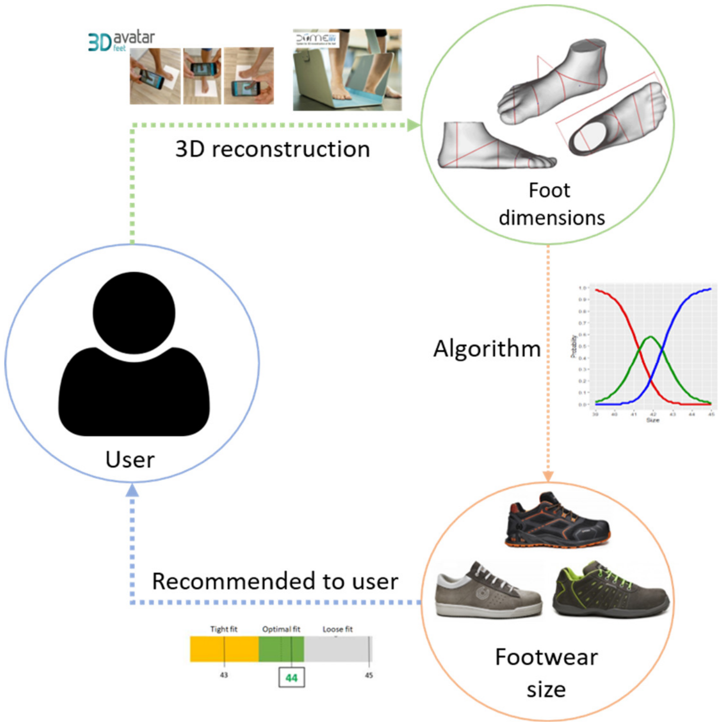

Figure 11. Overview of the size recommendation system.

3. Results and Discussion

3.1. Size recommendation

On average, the fit tolerance was 82%. It varied between 61 to 95% depending on the shoe model. The two models that showed a fit tolerance below 70% were models with fit issues that probably affected the size choices of the subjects. These results are in line with similar systems using logistic regressions to recommend sizes [19], [20] and provides better results than just picking the usual size (55-60%) or using the brand's size chat using foot length (45-55%).

Each shoe model of the system was trained just using fit trials from three consecutive sizes but it was tested using the full size span. Success rate was found similar for all the sizes. Figure 12 shows how the foot with failed recommendations (blue dots) are distributed along the full size span (red dots).

Figure 12. Failed recommendations (blue dots) on the IBV's database (red dots).

The feet that showed a lower performance of the system were examined through its features. They were geometrically similar to the feet which showed high performance. We think that the differences of the size choice could be due to differences in the individual fit preference of the subjects. Yet, the fit preference, as it was gathered in this work, was not included in the algorithm because it did not provide any improvement in the performance. This could be either because self-reported preference does not provide a reliable information about the actual fit preferences, or because the way we did it, with a simple binary scale was not suitable. The former is supported by the identification of some cases where the subjects' choices were not consistent with the reported preference.

3.2. Insole recommendation

On average, the percentage of fit tolerance is 92%. It varied between 88 to 100% depending on the insole type.

3.3. Market deployment

SCAN&FIT booth and app were released to market in 2020. In 2023, there are more than 180 points of sale equipped with SCAN&FIT booths around the world. Since its release, there have been more than 14K foot reconstructions made with the system in order to get different size recommendations, 70% made from the booth and 30% made from the phone app. This new technology is fast and highly usable and it is very valued by partners and customers, i.e. the retailers and the health and safety departments of large companies. Moreover, it provides a "living" source of data of the customers' feet, which will be used for future improvements of the products.

4. Conclusions, limitations and further research

Foot features are crucial for accurate size recommendations, emphasizing the need to ensure data quality. Collecting and maintaining high-quality foot measurements will enhance the reliability and precision of the size recommendation system.

This size recommendation system has demonstrated its effectiveness in providing accurate recommendations to users using logistic regression.

Furthermore, our approach showed similar results than previous projects but using less training data; namely, focusing on three consecutive sizes and using a projection technique instead of testing the full-size span. This efficiency provided a great advantage in the footwear case, where the size span is of 10+ sizes. It could also make a difference in products sold in a reduced number of sizes (S, M, L, XL) but have a high renewal rate and thus need a recursive effort to retrain the system (e.g. fast fashion).

The recommendation system's adaptability to different domains has been pivotal in its successful deployment. Its simple architecture allows and easy integration into different applications and platforms, making it a versatile and widely applicable solution. It also has low computation requirements and its output is very rich, which offers many possibilities to create interfaces that help the final consumers to make an informed decision about what size to choose.

However, further research is needed to address certain challenges and areas for improvement. The influence of personal preference is a significant factor in determining footwear size, contributing to around 20-30% of the variability. This highlights the importance of considering individual preferences when providing size recommendations.

Despite the training of the algorithms was more efficient than previous ones, the cold-start problem for new items still remains a concern. Further efforts are required to optimize the system's performance in such scenarios. Future work should focus on scaling the training process to enable the system to adapt to new shoe models without the need for costly experimentations.

Overall, the findings of this paper demonstrate the potential and significance of recommendation systems in enhancing user experience, driving engagement, and facilitating personalized content delivery in diverse applications and domains. The continued development and refinement of such systems hold promising prospects for the advancement of personalized user interactions and content recommendations in various digital platforms and services.

References

[1] A. K. Buldt and H. B. Menz, “Incorrectly fitted footwear, foot pain and foot disorders: a systematic search and narrative review of the literature,” *J. Foot Ankle Res.*, vol. 11, no. 1, p. 43, Jul. 2018, doi: 10.1186/s13047-018-0284-z.

[2] A. B. Dufour *et al.*, “Foot pain: Is current or past shoewear a factor?,” *Arthritis Care Res.*, vol. 61, no. 10, pp. 1352–1358, 2009, doi: 10.1002/art.24733.

[3] I. Russo and S. Cardinali, “Product Returns and Customer Value: A Footware Industry Case,” 2012, pp. 79–97.

[4] S. J. Marr and S. Quine, “Shoe concerns and foot problems of wearers of safety footwear,” *Occup. Med.*, vol. 43, no. 2, pp. 73–77, Jan. 1993, doi: 10.1093/occmed/43.2.73.

[5] R. Orr, D. Maupin, R. Palmer, E. F. D. Canetti, V. Simas, and B. Schram, “The Impact of Footwear on Occupational Task Performance and Musculoskeletal Injury Risk: A Scoping Review to Inform Tactical Footwear,” *Int. J. Environ. Res. Public. Health*, vol. 19, no. 17, p. 10703, Aug. 2022, doi: 10.3390/ijerph191710703.

[6] S. Singh, G. M. Abdulla, S. Borar, and S. Arora, “Footwear Size Recommendation System.” arXiv, Jun. 28, 2018. doi: 10.48550/arXiv.1806.11423.

[7] R. Guigourès, Y. K. Ho, E. Koriagin, A.-S. Sheikh, U. Bergmann, and R. Shirvany, “A hierarchical bayesian model for size recommendation in fashion,” in *Proceedings of the 12th ACM Conference on Recommender Systems*, in RecSys ’18. New York, NY, USA: Association for Computing Machinery, Sep. 2018, pp. 392–396. doi: 10.1145/3240323.3240388.

[8] A. Nestler, N. Karessli, K. Hajjar, R. Weffer, and R. Shirvany, “SizeFlags: Reducing Size and Fit Related Returns in Fashion E-Commerce,” in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, in KDD ’21. New York, NY, USA: Association for Computing Machinery, Aug. 2021, pp. 3432–3440. doi: 10.1145/3447548.3467160.

[9] A. Ballester *et al.*, “Fast, Portable and Low-Cost 3D Foot Digitizers: Validity and Reliability of Measurements,” presented at the Proc. of 3DBODY.TECH 2017 - 8th Int. Conference and Exhibition on 3D Body Scanning and Processing Technologies, Montreal QC, Canada, 11-12 Oct. 2017, Oct. 2017. doi: 10.15221/17.218.

[10] A. Jurca, J. Žabkar, and S. Džeroski, “Analysis of 1.2 million foot scans from North America, Europe and Asia,” *Sci. Rep.*, vol. 9, no. 1, p. 19155, Dec. 2019, doi: 10.1038/s41598-019-55432-z.

[11] A. Alcacer, I. Epifanio, J. Valero, and A. Ballester, “Combining Classification and User-Based Collaborative Filtering for Matching Footwear Size,” *Mathematics*, vol. 9, no. 7, Art. no. 7, Jan. 2021, doi: 10.3390/math9070771.

[12] S. Huang, Z. Wang, and Y. Jiang, “Guess your size: A hybrid model for footwear size recommendation,” *Adv. Eng. Inform.*, vol. 36, pp. 64–75, Apr. 2018, doi: 10.1016/j.aei.2018.02.003.

[13] D. Omrcen and A. Jurca, “Shoe Size Recommendation System Based on Shoe Inner Dimension Measurement,” in *Proceedings of the 2nd International Conference on 3D Body Scanning Technologies, Lugano, Switzerland, 25-26 October 2011*, Lugano, Switzerland: Hometrica Consulting - Dr. Nicola D’Apuzzo, Oct. 2011, pp. 158–163. doi: 10.15221/11.158.

[14] M. Costea, A. Mihai, and A. Seul, *3D Modelling Of Customized Lasts Based On Anthropometric Data Acquired From 3D Foot Scanning - One Study Case*. 2021. doi: 10.21203/rs.3.rs-394662/v1.

[15] A. Jurca and S. Dzeroski, “Foot length and width variety of shoppers who purchased the same style and size of running shoes,” *Footwear Sci.*, vol. 15, no. sup1, pp. S115–S116, Jun. 2023, doi: 10.1080/19424280.2023.2199390.

[16] E. Parrilla *et al.*, “Low-cost 3D foot scanner using a mobile app,” *Footwear Sci.*, vol. 7, no. sup1, pp. S26–S28, 2015.

[17] M. Kouchi *et al.*, “IEEE SA 3D Body Processing Industry Connections–Comprehensive Review of Foot Measurements Terminology in Use,” *IEEE SA 3D Body Process. Ind. Connect. Rev. Foot Meas. Terminol. Use*, pp. 1–63, Jan. 2021.

[18] J. García-Hernández *et al.*, “The MORFO3D Foot Database,” in *Pattern Recognition and Image Analysis*, J. S. Marques, N. Pérez de la Blanca, and P. Pina, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 658–665. doi: 10.1007/11492542_80.

[19] S. Alemany *et al.*, “Exploitation of 3D Body Databases for the Apparel Industry,” in *Proceedings of the 4th International Conference on 3D Body Scanning Technologies, Long Beach CA, USA, 19-20 November 2013*, Long Beach CA, USA: Hometrica Consulting - Dr. Nicola D’Apuzzo, Nov. 2013, pp. 456–466. doi: 10.15221/13.456.

[20] A. Ballester *et al.*, “Kidsize: Always Get the Right Size,” presented at the Proc. of 7th Int. Conference on 3D Body Scanning Technologies, Lugano, Switzerland, Nov. 2016, Nov. 2016. doi: 10.15221/16.139.