# FINAL REPORT

**SYS660 – Decision and risk analysis**
**Group 7- Weiwei Tao, Priya Rao**
**Topic – Shoe Recommendation System**

**Problem Statement**

The main goal of the shoe recommendation system is to provide extremely precise and customized footwear recommendations by utilizing contextual data and user-specific preferences. A software that recommends shoes to consumers based on factors like price, gender, usage, material, foot length, customers, waterproofness, and comfort level is called a shoe recommendation system. Retail websites and e-commerce platforms that sell shoes often use this specific type of recommendation system. The system uses sophisticated algorithms to examine each user's unique profile, which includes demographic information, preferred styles, previous purchases, and behavior, in addition to a thorough shoe inventory, with the goal of improving the overall user experience. The objective is to produce personalized recommendations that take into account the price, gender, brand, material, feet length, customer reviews, waterproof level, comfortability level and budgetary restrictions of each user, while also catering to their own interests, wants, and situations.

**Development Procedure**

In order to build a useful and customized platform, there are a few essential elements in the development process of a shoe suggestion system. First, extensive systems for collecting data are set up to obtain a variety of user data, such as demographics, preferred styles, and past purchases. At the same time, a comprehensive shoe inventory is created, which includes several aspects such as price, brand, usage, material, feetlength, customer reviews, waterproof level and comfortability level. The gathered data is then refined and made ready for analysis using data preparation procedures. In order to extract pertinent shoe and user attributes and produce recommendations that are correct, feature engineering and selection techniques are used.

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | Price | Gender | Brand | Usage | Material | Feet Lengt | Customer | Waterproc | Comfortability Level | |
| 2 | 0 | 70.98 | Unisex | Puma | Running | Synthetic | 9 | 4.5 | Medium | Low | |
| 3 | 1 | 156.61 | Male | Under Arm | Hiking | Leather | 11 | 4.9 | Low | Low | |
| 4 | 2 | 172.56 | Unisex | New Balar | Casual | Suede | 7 | 5 | Medium | High | |
| 5 | 3 | 129.61 | Unisex | New Balar | Hiking | Leather | 10 | 4.5 | High | High | |
| 6 | 4 | 88.83 | Female | Adidas | Running | Suede | 6 | 1.6 | Low | Medium | |
| 7 | 5 | 103.7 | Female | Nike | Casual | Synthetic | 5 | 4.9 | Medium | Low | |
| 8 | 6 | 70.44 | Female | New Balar | Walking | Canvas | 9 | 3.2 | High | Medium | |
| 9 | 7 | 120.88 | Female | Asics | Running | Mesh | 7 | 3.4 | Low | High | |
| 10 | 8 | 141.45 | Male | Skechers | Training | Mesh | 5 | 1.4 | Low | Medium | |
| 11 | 9 | 187.74 | Male | Skechers | Training | Synthetic | 13 | 1 | Low | High | |
| 12 | 10 | 58.99 | Unisex | Puma | Training | Mesh | 12 | 1.1 | Medium | High | |
| 13 | 11 | 158.85 | Male | Reebok | Hiking | Leather | 13 | 4.7 | High | Low | |
| 14 | 12 | 91.87 | Unisex | Reebok | Running | Synthetic | 10 | 3.8 | Low | Medium | |
| 15 | 13 | 116.89 | Female | Reebok | Running | Leather | 5 | 2.4 | Medium | Low | |
| 16 | 14 | 79.18 | Unisex | New Balar | Casual | Suede | 6 | 4.2 | High | Low | |
| 17 | 15 | 127.73 | Female | Reebok | Training | Canvas | 7 | 1.7 | Medium | Low | |
| 18 | 16 | 154.47 | Female | Nike | Training | Mesh | 6 | 3.5 | Medium | High | |
| 19 | 17 | 192.21 | Female | Vans | Hiking | Synthetic | 5 | 3.8 | Medium | Medium | |
| 20 | 18 | 153.42 | Female | Under Arm | Training | Synthetic | 11 | 3.2 | High | High | |
| 21 | 19 | 112.35 | Female | Adidas | Running | Mesh | 8 | 3.3 | Low | Medium | |
| 22 | 20 | 173.87 | Female | Asics | Walking | Synthetic | 5 | 4.1 | Low | Low | |
| 23 | 21 | 78.78 | Female | Skechers | Training | Synthetic | 8 | 2.3 | Low | High | |
| 24 | 22 | 107.91 | Male | Under Arm | Training | Mesh | 10 | 2.1 | Medium | High | |
| 25 | 23 | 132.43 | Unisex | Converse | Training | Leather | 9 | 4.1 | Medium | Medium | |
| 26 | 24 | 77.9 | Female | New Balar | Casual | Suede | 6 | 4.7 | High | High | |
| 27 | 25 | 107.98 | Female | Puma | Training | Canvas | 12 | 3.8 | Medium | Low | |

Concurrently, the experience design and user interface are developed iteratively with the goal of producing a platform that is simple to use and intuitive. The UI makes it simple for customers to enter their preferences and display suggested shoes with full product details, photos, costs, and links to make purchases. By adding feedback systems, consumers can offer their opinions on recommended shoes, which helps to further improve the system's suggestions.Ensuring data privacy and security safeguards, protecting user information, is prioritized throughout the development process. The system is updated often to incorporate newly released shoes, modify algorithms in response to user comments, and keep up with evolving fashion trends.

Performance testing and scalability are essential to ensuring the system can handle expanding user traffic and shoe inventory while still being accurate and responsive. The development process consists of data collection, preprocessing, algorithm implementation, interface design, security protocols, and continuous updates. The goal is to build a reliable shoe recommendation system that increases user engagement and satisfaction by providing tailored and accurate suggestions.

**Working and analysis of the shoe recommendation system**

**Dataset Description**

Some of the attributes chosen from the dataset are price, gender, brand, usage, material, feet length, customer reviews, waterproof level, comfortability level.The dataset is described below:

- Price -The cost of the shoes is represented by the price attribute.It might be a number that represents the price of the shoes, broken down into various price points or ranges.
- Gender – The target gender of the shoes is indicated by this characteristic. It may consist of classifications like kids' shoes, women's shoes, or men's shoes, or it may be unisex, male, or female.
- Brand - This characteristic speaks to the shoe's maker or brand. It might feature a variety of brands, like New Balance, Adidas, Puma, Reebok, and Nike.
- Usage - This feature explains the intended function or purpose of the shoes. It may cover things like sports, hiking, work, casual and formal attire, running, etc.
- Material - The main component that went into making the shoes is described in this feature. Materials such as suede, canvas, rubber, mesh, leather, and synthetics could be used
- Feet Length - The length of the foot that the shoes are appropriate for is represented by this characteristic. A number representing a shoe size or a measurement of the length of the foot could be displayed.
- Customer Reviews - This attribute includes customer feedback or reviews about the shoes. It might contain text data, ratings (e.g., star ratings), sentiment analysis, or aggregated scores based on customer reviews.
- Waterproof level - This feature indicates how water-resistant or waterproof the shoes are. It might have classifications like waterproof level to be low, medium and high.
- Comfortability level - This feature shows how comfy the shoes are. The classification of shoes into high, medium, and low comfort levels may be based on consumer feedback, product specifications, or specific comfort-related qualities.
- 

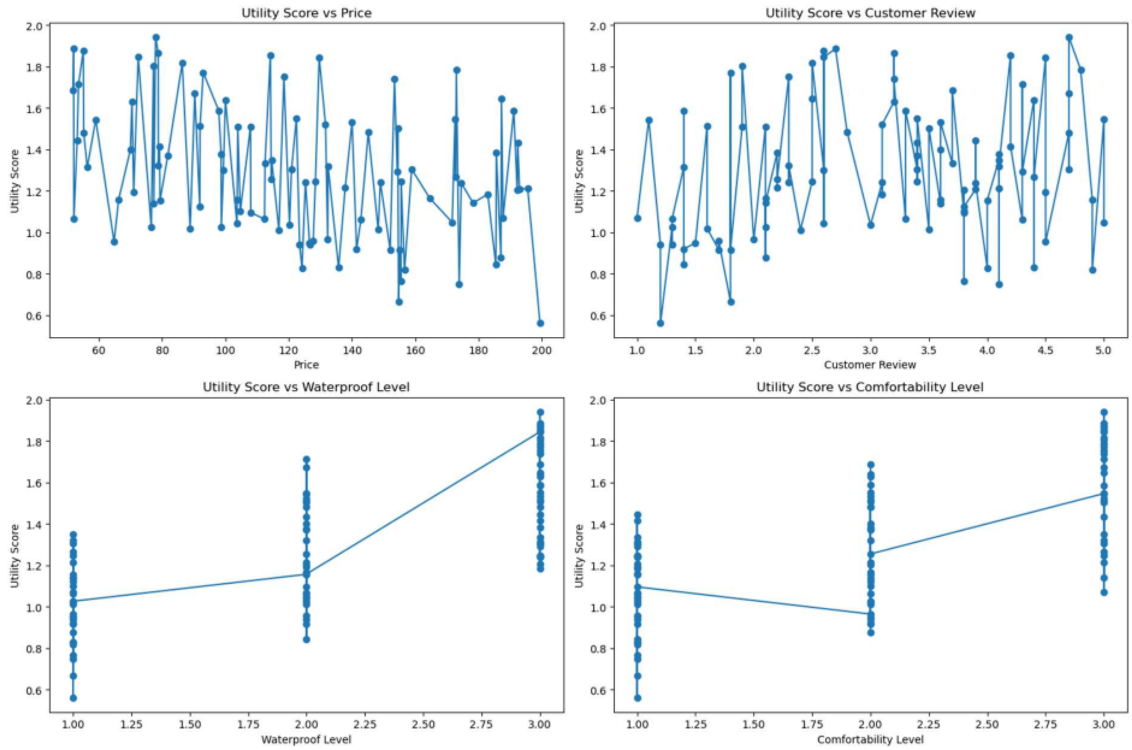
**Selection Criteria**

- Price
- Usage
- Customer Review
- Waterproof Level
- Comfortability Level

**Utility Function**

$$U = comfortability * weights_{comfortability} + review * weights_{review} + waterproof * weights_{waterproof} + (1 - price) * weights_{price}$$

**Utility Curve**




Utility Score vs Price:

The graph shows a general trend of decreasing utility score with increasing price. This indicates that, within the dataset, more expensive shoe models tend to have a lower utility score. This trend is consistent with the utility function where higher prices negatively impact the utility. The utility score appears to be relatively high for lower-priced shoes and decreases as the price increases, reflecting the significant impact of price on the overall utility. This suggests that consumers might favor affordability in their shoe choices, or that lower-priced shoes in the dataset happen to align better with other desirable features.

Utility Score vs Customer Review:

The utility score generally increases with higher customer reviews, demonstrating a positive correlation between customer satisfaction and the utility score. This graph illustrates that shoes with higher customer reviews are perceived as more desirable or useful, as per the utility function. The trend is mostly linear, indicating a steady increase in utility with better reviews. This trend emphasizes the importance of customer satisfaction in determining the overall desirability of a product.

Utility Score vs Waterproof Level:

This graph shows a step-like increase in utility scores with higher waterproof levels. Shoes categorized as high in waterproofing tend to have a significantly higher utility score compared to medium or low waterproof levels. This trend indicates that waterproofing is an important factor for utility, and shoes with better waterproofing capabilities are considerably more desirable in the dataset. The discrete levels of waterproofing create distinct groups in terms of utility, underscoring the value placed on this feature by consumers.

Utility Score vs Comfortability Level:
Similar to the waterproof level, the utility score increases with higher comfortability levels. The graph depicts a clear preference for shoes with higher comfortability, as evidenced by the marked increase in utility scores from low to medium, and medium to high comfortability levels. This suggests that comfort is a significant factor in determining the desirability of shoes, with highly comfortable shoes being much more preferred, according to the utility model used. The distinct jumps in utility between comfort levels highlight the substantial impact of comfort on consumer preferences.

**System Logics**

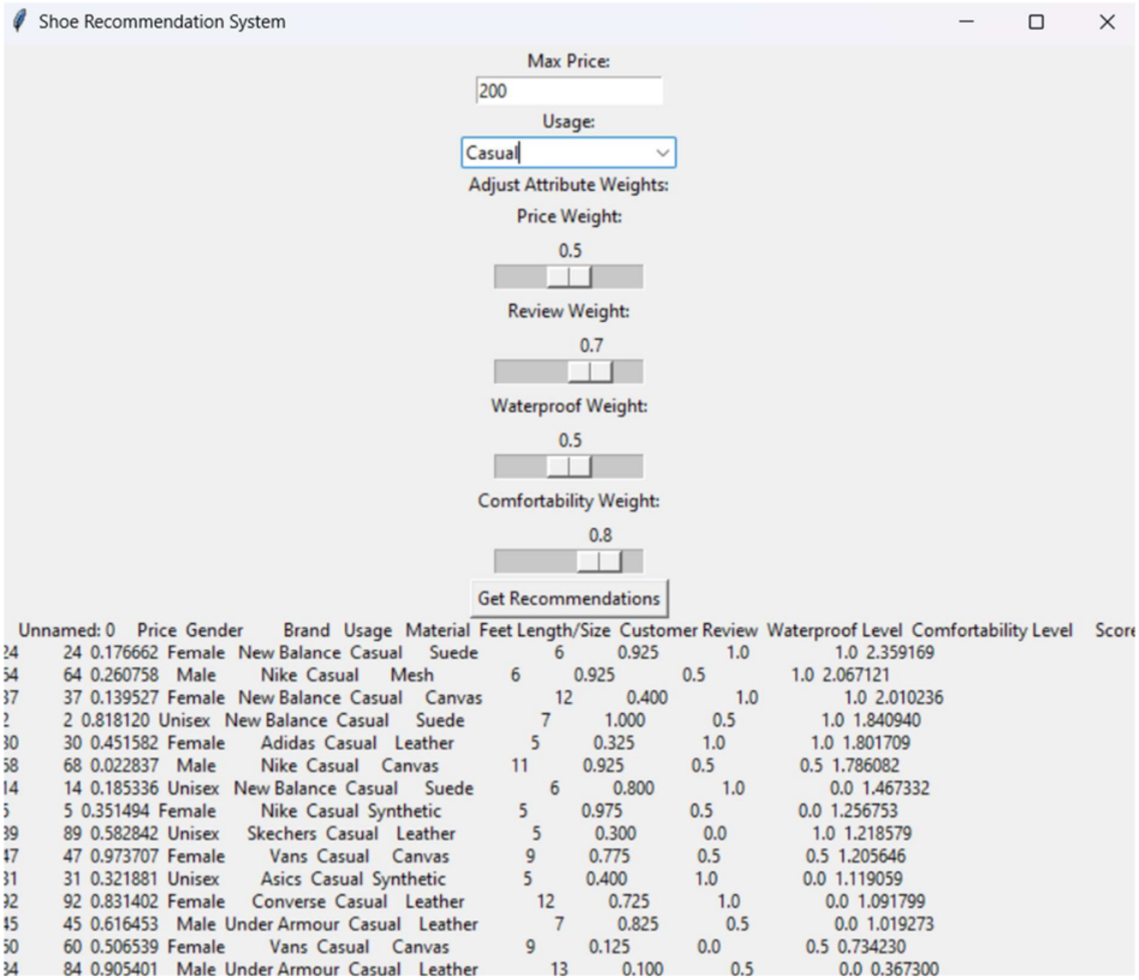
In our Python code, the shoe recommendation function, ‘get_recommendations’, employs a Monte Carlo simulation approach to provide tailored shoe recommendations based on user preferences. The function first filters the dataset (df) for shoes within a specified maximum price (max_price) and intended usage (usage). It then normalizes the values in the columns 'Price', 'Customer Review', 'Waterproof Level', and 'Comfortability Level' to ensure comparability.

The core of the function lies in its Monte Carlo simulation, where it runs ‘num_simulations’ (defaulted to 1000) iterations. In each iteration, it slightly varies the weights assigned to each attribute (price, customer review, waterproof level, comfortability level) within a specified range (±10% of the original weights). This introduces randomness to simulate real-world variations in consumer preferences. The utility score for each shoe is calculated in every simulation, factoring in these randomly adjusted weights.

After all simulations, the function averages the utility scores for each shoe to get a 'Monte Carlo Score', which is less sensitive to the specific weights and more robust to variations in customer preferences. Finally, it sorts the shoes based on this score, with higher scores indicating more recommended options. This method provides a comprehensive and adaptive recommendation that accounts for a range of consumer preferences, making it a versatile tool for personalized shoe selection.

**System Interface**

The shoe recommendation system interface was developed using Python programming language.This recommendation system has two input columns named maximum price and kind of usage with four attribute columns as price weight, review weight, waterproof weight, comfortability weight. First, the user/customer inputs the maximum price as $200 , the usage of the shoe is chosen from the dropdown as Casual. The four available usage options are casual, running, hiking and walking.The third step is to adjust the attribute weights like price weights, review weights, waterproof weights and comfortability weights.The customer has to adjust the price weights to 0.5, review weights to 0.7, waterproof weights to 0.5 and comfortability weights to 0.8 respectively. Finally click the Get Recommendations option below to generate the necessary recommendations.There would be numerous shoe recommendations displayed on the screen with different prices, brands, gender(female/male/unisex), usage, material, FeetLength/size, representing different customer reviews, waterproof level and comfortability level.


Shoe Recommendation System
Max Price:
200
Usage:
Casual
Adjust Attribute Weights:
Price Weight:
0.5
Review Weight:
0.7
Waterproof Weight:
0.5
Comfortability Weight:
0.8
Get Recommendations
Unnamed: 0 Price Gender Brand Usage Material Feet Length/Size Customer Review Waterproof Level Comfortability Level Scor
24 0.176662 Female New Balance Casual Suede 6 0.925 1.0 1.0 2.359169
64 0.260758 Male Nike Casual Mesh 6 0.925 0.5 1.0 2.067121
37 0.139527 Female New Balance Casual Canvas 12 0.400 1.0 1.0 2.010236
2 0.818120 Unisex New Balance Casual Suede 7 1.000 0.5 1.0 1.840940
30 0.451582 Female Adidas Casual Leather 5 0.325 1.0 1.0 1.801709
68 0.022837 Male Nike Casual Canvas 11 0.925 0.5 0.5 1.786082
14 0.185336 Unisex New Balance Casual Suede 6 0.800 1.0 0.0 1.467332
5 0.351494 Female Nike Casual Synthetic 5 0.975 0.5 0.0 1.256753
89 0.582842 Unisex Skechers Casual Leather 5 0.300 0.0 1.0 1.218579
47 0.973707 Female Vans Casual Canvas 9 0.775 0.5 0.5 1.205646
31 0.321881 Unisex Asics Casual Synthetic 5 0.400 1.0 0.0 1.119059
92 0.831402 Female Converse Casual Leather 12 0.725 1.0 0.0 1.091799
45 0.616453 Male Under Armour Casual Leather 7 0.825 0.5 0.0 1.019273
60 0.506539 Female Vans Casual Canvas 9 0.125 0.0 0.5 0.734230
84 0.905401 Male Under Armour Casual Leather 13 0.100 0.5 0.0 0.367300