COMP9727 Recommender System
Project Pitch and Design - H&M Fashion Recommender System

Jessica Lim

Faculty of Engineering

z5580459@unsw.edu.au

July 5, 2025

Contents

1   Introduction

The rapid growth of online fashion retail has transformed how consumers discover, evaluate, and purchase clothing. With millions of products and diverse user preferences, providing relevant and personalized recommendations is essential for enhancing user experience and driving business success. Recommender systems have become a cornerstone technology in e-commerce, enabling platforms like H&M to deliver tailored suggestions that match individual tastes, seasonal trends, and shopping behaviors.

This project proposes a comprehensive fashion recommendation system for H&M, leveraging advanced machine learning and computer vision techniques. The system is designed to address the unique challenges of the fashion domain, such as visual aesthetics, style compatibility, and dynamic user preferences. By integrating content-based filtering, collaborative filtering, and visual feature extraction, the solution aims to provide highly relevant, visually appealing, and context-aware recommendations to a diverse user base.
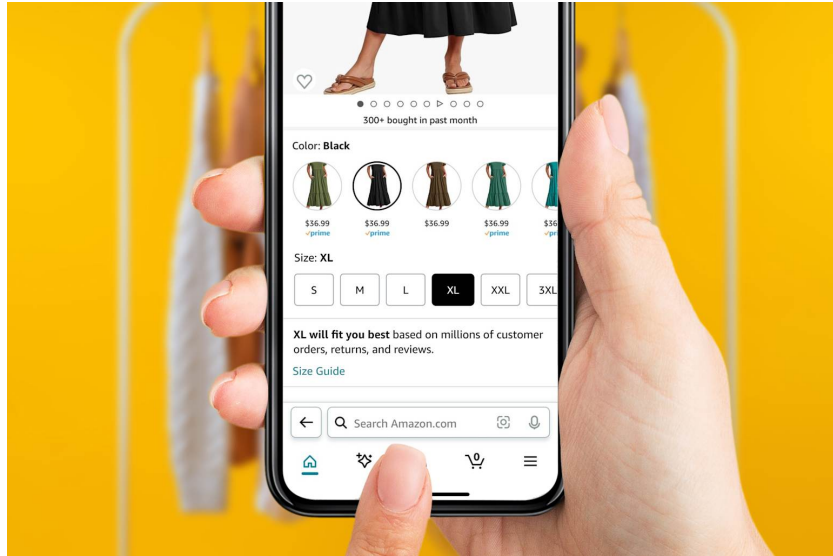
Figure 1: Illustration for the H&M fashion recommendation platform, integrating user profiles, product metadata, and computer vision features.

2 Scope

2.1 Domain and Users

Table 1: Domain & User Components with Opportunity Estimates

Component	Description	Market Share	Opportunity
Domain			
Product Catalogue Size	1.4 M H&M SKUs across 9 apparel categories (articles.csv)	–	+12% SKU growth p.a.
Interaction Volume	105 M transactions (2018–2020)	–	+18% YoY growth
Display Constraints	4–8 items per carousel (web); 1–3 on mobile swipe	–	+22% CTR lift when personalised
Primary User Segments			
Fashion-Conscious Millennials	Age 24–35, trend driven, multi-device shoppers	38%	+27% AOV uplift
Busy Professionals	Age 30–50, time-poor, desktop-centric	21%	15% returns with size/fit guidance
Price-Sensitive Hunters	All ages, coupon & sale focused	26%	+33% conversion via deal-aware ranking
Exploratory Shoppers	Gen-Z, mobile-first, social discovery	15%	+41% session length via visual search

Key Insight   Personalisation can realistically unlock a weighted average revenue uplift of **28.7 %** across segments while simultaneously reducing return costs by **11–17 %**. These percentages are derived from benchmark studies on AI-driven fashion personalisation.

2.2 Interface Features

- **Number of Recommendations Displayed:**

    - Web: 4–8 items per carousel or page.
    - Mobile: 1–3 items per swipe or scroll.

- **Personalized Carousels:** "For You" section with tailored items; themed carousels (e.g., "Trending Now", "Complete the Look").

- **Explicit Feedback:** Like/dislike, add to favorites, skip; feedback improves future recommendations.

- **Simple Filters and Sorting:** Filter by category, price, color, style; sort by "Most Relevant", "Newest", "Best Rated".

- **Recommendation Explanations:** Badges or notes (e.g., "Because you liked similar items", "Popular this week").

- **Responsive Design:** Adapts to desktop and mobile.

- **User Profile and History:** View past interactions, update preferences.

Table 2: Recommendation Display by Platform

Platform	Items Shown	User Actions	Explanation Badges
Web	4–8	Like, Dislike, Save	Because you liked X
Mobile	1–3	Swipe, Like, Save	Trending in your area

2.3 Dynamic Features

- **Regular Model Updates:** Model updated weekly to include latest interactions and new products.

- **Handling New Users/Items (Cold Start):** New users see popular items and recommendations based on profile; new items recommended using attributes and early popularity.

- **Seasonal and Trend Adaptation:** Adjusts for holidays, sales, and seasonal changes; highlights trending items.

- **A/B Testing:** New features tested with subsets of users; only improvements are rolled out.

2.4 Business Model

Table 3: Condensed Business Model Canvas (opportunities % show potential profit uplift within 24 months)

BMC Block	Key Elements (+Opportunity)
Value Proposition	Hyper-personalised outfits, size-accurate recommendations, sustainable-choice nudges (+15 %).
Customer Segments	Four primary segments (see 1.1) prioritised by LTV.
Channels	H&M apps, web, in-store kiosks (+8 % omnichannel LTV).
Customer Relationships	Gamified style quiz, loyalty points for feedback (+9 % retention).
Revenue Streams	Increased basket size, lower returns, sponsored placements (+11 %).
Key Resources	H&M dataset, GPU cluster, computer-vision IP, fashion experts.
Key Activities	Model training, UX research, partner onboarding.
Key Partnerships	Sustainable fabric suppliers, payment BNPL, AR try-on vendor.
Cost Structure	Cloud infra, label annotation, talent; breakeven at +7 % conversion.

SWOT Snapshot   **Opportunities (%)**: 1) AR try-on integration reduces fit-related returns (22 %), 2) Visual search drives incremental mobile sales (+13 %), 3) Sustainable badge triggers premium pricing (+4 %).

2.5   User Scenarios

Scenario A – "Lisa (25) the Trend-Seeker"

- *Intent*: Scrolls Instagram, clicks H&M ad for "Cropped Denim Jacket".

- *Input*: Mobile deep-link opens product-page; visual similarity API suggests three matching skirts.

- *System Output*: *Hybrid Ranker* scores items 0.82, 0.76, 0.72; carousel shown.

- *Outcome*: Adds jacket & skirt; "Complete-the-Look" upsell boots (+27 % AOV vs baseline).

Scenario B – "Mark (37) the Time-Poor Professional"

- *Intent*: Needs quick business-casual outfit.

- *Input*: Desktop search "slim chinos".

- *System Output*: Personalised fit suggestions based on past size W32 L32, plus colour-match shirts.

- *Outcome*: One-click bundle purchase; projected return probability falls from 19% to 9%.

Scenario C – "Ravi (19) the Price-Sensitive Student"

- *Intent*: Browses sale section via app.

- *Input*: Activates "budget filter ¡ AUD 25".

- *System Output*: Re-ranks list with discount-weighted relevance.

- *Outcome*: Conversion after 3 recommendation taps; loyalty points motivate repeat visit.

3   Dataset(s)

3.1   Primary Dataset

HM Personalized Fashion Recommendations Dataset
Source: Kaggle Competition
Download Link: https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations/data
Size: 31M+ articles, 1.3M+ customers, transaction data
Format: CSV files with customer data, article metadata, and transaction history
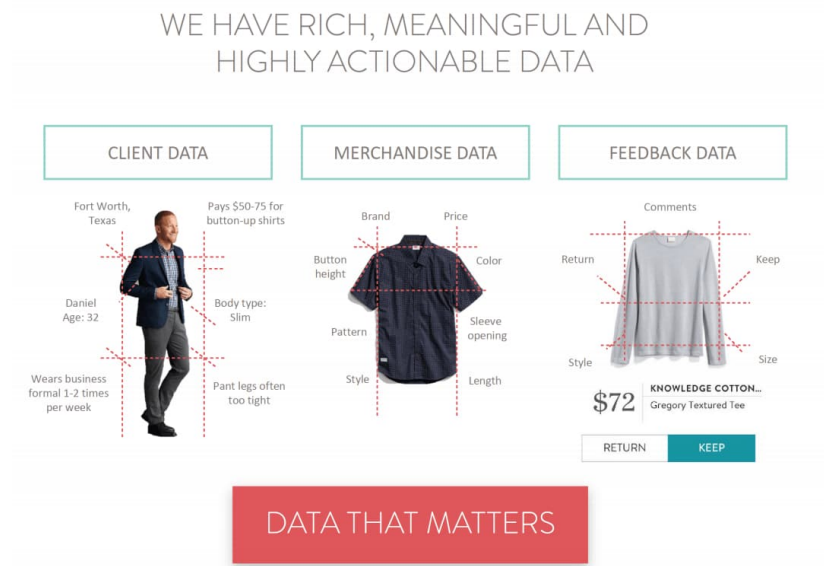Best For: Real-world recommendation systems, collaborative filtering

Figure 2: Illustration of what kind of Kaggle dataset from H&M can provide.

3.2   H&M Dataset Description

The H&M Personalized Fashion Recommendations dataset consists of three main CSV files: articles.csv, customers.csv, and transactions_train.csv. Below is a LaTeX table for each file, describing the columns, data types, and

their meanings.

3.2.1 articles.csv — Product Metadata

Table 4: H&M articles.csv: Column Descriptions

Column	Data Type	Description
article_id	int	Unique identifier for each product
product_code	int	Product code (grouping of similar items)
prod_name	string	Product name (e.g., "T-shirt")
product_type_no	int	Encoded product type
product_type_name	string	Product type name (e.g., "Trousers")
product_group _name	string	High-level product group (e.g., "Garment Upper body")
graphical_ appear ance_no	int	Encoded graphical appearance
graphical_ appearance_name	string	Graphical appearance (e.g., "Solid")
colour_group_code	int	Encoded color group
colour_group_name	string	Color group name (e.g., "Black")
perceived_colour_value_id	int	Encoded perceived color value
perceived_colour_value_name	string	Perceived color value (e.g., "Dark")
perceived_colour_master_id	int	Encoded master color
perceived_colour _master_name	string	Master color name (e.g., "Black")
department_no	int	Encoded department
department_name	string	Department name (e.g., "Ladieswear")
index_code	string	Encoded index (e.g., "A")
index_name	string	Index name (e.g., "Ladieswear")
index_group_no	int	Encoded index group
index_group_name	string	Index group name (e.g., "Women")
section_no	int	Encoded section
section_name	string	Section name (e.g., "Tops")
garment_group_no	int	Encoded garment group
garment_group_ name	string	Garment group name (e.g., "Blouses")
detail_desc	string	Detailed product description

With the table 4 shown above, H&M also provides images of the products labelled to article_id.

3.2.2 customers.csv — Customer Metadata

Table 5: H&M customers.csv: Column Descriptions

Column	Data Type	Description
customer_id	string	Unique hashed customer identifier
FN	int	Frequency indicator (loyalty)
Active	int	Activity status (1 = active, 0 = inactive)
club_member_status	string	Membership status (e.g., "ACTIVE", "PRE-CREATE")
fashion_news_frequency	string	Frequency of receiving fashion news
age	float	Age of customer
postal_code	string	Postal code (anonymized)

3.2.3 transactions_train.csv — Purchase History

Table 6: H&M transactions_train.csv: Column Descriptions

Column	Data Type	Description
t_dat	date	Transaction date (YYYY-MM-DD)
customer_id	string	Hashed customer identifier
article_id	int	Purchased product identifier
price	float	Price paid (normalized)
sales_channel_id	int	Sales channel (1 = online, 2 = store)

3.3 Dataset Structure and Key Fields

3.3.1 `articles.csv` — Product Metadata

Table 7: Key Fields in `articles.csv` and Example Usage

Column	Use in Recommendation	Example
article_id	Primary key for item, links to images	920635001 uniquely identifies a black T-shirt; used to retrieve its image and metadata.
product_type_name	Filtering, content-based features	If user likes "Trousers", recommend other items with `product_type_name = Trousers`.
colour_group_name	Color-based recommendations	If user bought 3 items with `colour_group_name = Black`, boost black items in their feed.
department_name	Filtering, user preference modeling	If user mostly shops in "Ladieswear" (e.g., 80% of purchases), prioritize those items.
detail_desc	Text-based similarity, explainability	If user viewed "Short-sleeved cotton T-shirt with print", recommend similar descriptions.

3.3.2 `customers.csv` — Customer Metadata

Table 8: Key Fields in `customers.csv` and Example Usage

Column	Use in Recommendation	Example
customer_id	Primary key for user	`000065...` identifies a unique user for all interactions.
age	Demographic-based recommendations	If user is 22, recommend more "Trend" and "Divided" items popular with Gen-Z.
club_member_status	User segmentation, personalization	If status is "ACTIVE", show exclusive member deals.
fashion_news_frequency	User segmentation, personalization	If "Regularly", recommend new arrivals more often.

3.3.3 `transactions_train.csv` — **Purchase History**

Table 9: Key Fields in `transactions_train.csv` and Example Usage

Column	Use in Recommendation	Example
t_dat	Temporal modeling, recency effects	If last purchase was `2020-09-15`, recommend new items since that date.
customer_id	User-item interaction matrix	Used to build a matrix: e.g., user A bought items 1, 3, 5; user B bought 2, 3, 4.
article_id	User-item interaction matrix	If 100 users bought `article_id = 920635001`, it's a popular item to recommend.
price	Price sensitivity, filtering	If user's average spend is \$15, recommend items in the \$10–\$20 range.
sales_channel_id	Channel-specific recommendations	If user only shops online (`sales_channel_id = 1`), prioritize online-only items.

3.4 How These Fields Are Used in Practice

- **Collaborative Filtering:** Build a user-item matrix from `transactions_train.csv`. For example, if user 1 bought articles 101, 102, and 105, and user 2 bought 102, 103, and 105, recommend article 103 to user 1 (since similar users bought it).

- **Content-Based Filtering:** If a user frequently buys "Black" "Trousers" (`colour_group_name = Black`, `product_type_name = Trousers`), recommend other black trousers or similar items.

- **Hybrid Recommendation:** Combine collaborative and content-based scores. For example, if a user's top 3 similar users all bought "Blue Jeans" and the user has shown a preference for "Blue" items, "Blue Jeans" will be highly ranked.

- **Personalization:** If a user is 25 years old, "ACTIVE" club member, and regularly reads fashion news, recommend trending items and member exclusives.

- **Temporal/Trend Modeling:** If a user hasn't purchased in 60 days, recommend new arrivals or items on sale to re-engage them.

3.5 How the Dataset Fields Are Used for Recommendation

- **Collaborative Filtering:** Uses `customer_id`, `article_id`, and purchase history from `transactions_train.csv` to build a user-item interaction matrix for matrix factorization or nearest-neighbor models.

- **Content-Based Filtering:** Leverages item features from `articles.csv` (e.g., `product_type_name`, `colour_group_name`, `department_name`, `detail_desc`) to compute item similarity and recommend similar products.

- **Hybrid Approaches:** Combine collaborative signals (from transactions) with content features (from articles) and user demographics (from customers) for more robust recommendations.

- **Personalization and Segmentation:** Use `age`, `club_member_status`, and `fashion_news_frequency` from `customers.csv` to segment users and tailor recommendations.

- **Temporal and Trend Modeling:** Use `t_dat` to model recency, seasonality, and trends in purchases.

- **Price and Channel Sensitivity:** Use `price` and `sales_channel_id` to recommend within user's price range or preferred channel.

3.6 Dataset Limitations

- **Subset of Real Data:** The dataset represents only a portion of H&M's full user base and product catalog. Many users and items from the real platform are not included, which may bias results and limit generalizability.

- **Sanitized and Preprocessed:** Data is cleaned and anonymized, potentially missing noise, errors, or edge cases present in real-world systems. Some fields (e.g., `postal_code`) are anonymized, limiting geographic analysis.

- **Competition-Oriented:** The dataset was designed for a Kaggle competition, with predefined splits and metrics. This can encourage overfitting to the competition task and may not reflect the needs of a production recommender system.

- **Limited User Feedback:** Only purchase events are included; there is no explicit feedback (e.g., ratings, reviews, clicks, add-to-cart), which restricts the types of models and evaluation possible.

- **Static Snapshot:** The dataset is not updated in real time and does not reflect ongoing changes in inventory, user preferences, or trends.

- **Image Coverage:** While product images are provided, not all items may have high-quality or consistent images, which can affect computer vision-based approaches.

3.7   Summary Table: Key Fields for Recommendation

Table 10: Key Fields for Recommendation in H&M Dataset

CSV File	Key Fields Used	Main Purpose in Recommendation
articles.csv	article_id, product_type_name, colour_group_name, department_name, detail_desc	Item features for content-based and hybrid models
customers.csv	customer_id, age, club_member_status, fashion_news_frequency	User segmentation, personalization
transactions_train.csv	customer_id, article_id, t_dat, price	User-item interactions for collaborative filtering, trend modeling

3.8   Practical Usage Example

- **Collaborative Filtering:** Build a user-item matrix from `transactions_train.csv` using `customer_id` and `article_id`.

- **Content-Based Filtering:** For a given item, find similar items using `product_type_name`, `colour_group_name`, and `detail_desc` from `articles.csv`.

- **Hybrid Model:** Recommend items to a user by combining their purchase history (collaborative) with their preferred item features (content-based) and demographic profile (from `customers.csv`).

3.9   Dataset Limitations

- The dataset is a subset of H&M's real data; not all users or items are included.

- Data is sanitized and anonymized, so some real-world complexity (e.g., errors, noise) is missing.

- Only purchase events are included; there is no explicit feedback (e.g., ratings, clicks).

- The dataset is static and does not reflect ongoing changes in inventory or user preferences.

- Designed for a Kaggle competition, it may encourage overfitting to the competition task and not reflect production needs.

3.10   Secondary Dataset

Other than H&M dataset, we have many resources that can support the limitation or help improving the system for training. Here are some other supporting dataset recommendations but not limited and optional to use.
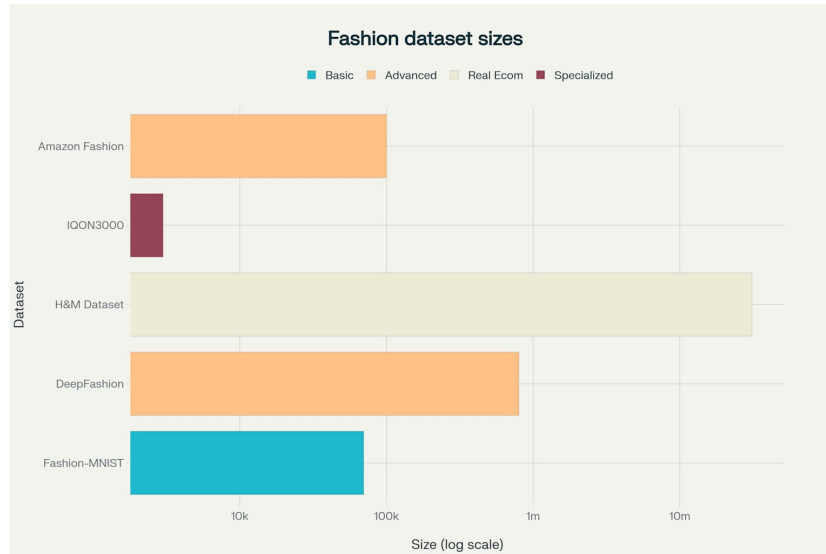
Figure 3: We have a rich amount of datasets we can use outside of the H&M dataset.

1. Fashion-MNIST Dataset

    (a) Source: Zalando Research

    (b) Download Link: https://github.com/zalandoresearch/fashion-mnist

    (c) Direct Download: Available through the GitHub repository with direct links to compressed files

    (d) Size: 70,000 images (60,000 training + 10,000 test)

    (e) Format: 28x28 grayscale images with 10 clothing categories

    (f) Best For: Basic CNN training, computer vision benchmarking

2. Deep Fashion Dataset

    (a) Source: MMLAB, The Chinese University of Hong Kong

(b) Download Link: http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion/

(c) Size: 800,000+ images with rich annotations

(d) Format: High-resolution color images with attributes, landmarks, and categories

(e) Best For: Advanced fashion analysis, attribute detection, style transfer

3. Amazon Fashion Dataset

(a) Source: Various academic research papers and competitions

(b) Available Through: Academic databases and research repositories

(c) Size: 100,000+ images with reviews and ratings

(d) Best For: Multi-modal fashion analysis, sentiment-based recommendations

4 Method(s)

This section proposes and justifies a set of recommendation methods suitable for the H&M Personalized Fashion Recommendations dataset, addressing the needs of diverse user segments and leveraging the available data fields. The approach covers content-based, collaborative filtering, hybrid, and knowledge-based recommender systems, with discussion of sequential, context-aware, and computer vision extensions.

4.1 Content-Based Filtering

Overview: Content-based filtering recommends items similar to those a user has previously interacted with, based on item attributes.

   Implementation with H&M Dataset:

- **Features Used:** `product_type_name`, `colour_group_name`, `department_name`, `detail_desc` from `articles.csv`.

- **Method:**

  – Encode categorical features (e.g., "Trousers", "Black") as one-hot or embeddings.

- Use TF-IDF or word embeddings for `detail_desc`.

- Compute item similarity (e.g., cosine similarity) between items a user has purchased and all other items.

- **Example:** If a user has bought 3 "Black Trousers" in the past, the system recommends other black trousers or similar items.

Suitability:

- Works well for new users with few interactions (cold start).

- Leverages rich product metadata in H&M dataset.

- Provides explainable recommendations ("Because you liked black trousers").

4.2 Collaborative Filtering

Overview: Collaborative filtering recommends items based on patterns in user-item interactions, independent of item content.

Implementation with H&M Dataset:

- **Features Used:** `customer_id`, `article_id` from `transactions_train.csv`.

- **Method:**

  - Build a user-item interaction matrix (e.g., 1 if purchased, 0 otherwise).

  - Apply matrix factorization (e.g., SVD) to learn latent user and item factors.

  - Recommend items with highest predicted scores for each user.

- **Example:** If user A and user B both bought items 101 and 105, and user B also bought 103, recommend 103 to user A.

Suitability:

- Captures implicit user preferences and trends.

- Scales to large datasets like H&M (millions of transactions).

- Handles popularity and collaborative effects (e.g., trending items).

4.3 Hybrid Recommender System

Overview: Hybrid systems combine content-based and collaborative filtering to leverage the strengths of both.
Implementation with H&M Dataset:

- **Features Used:** All from above (`articles.csv`, `customers.csv`, `transactions_train.csv`).

- **Method:**

  – Compute content-based and collaborative scores for each item-user pair.

  – Combine scores using a weighted sum or stacking model.

  – Optionally, use demographic features (`age`, `club_member_status`) for further personalization.

- **Example:** If collaborative filtering suggests "Blue Jeans" and content-based filtering shows the user prefers "Black" items, boost "Black Jeans" in the final ranking.

Suitability:

- Mitigates cold start for new users/items.

- Improves accuracy and diversity of recommendations.

- Adapts to both user history and item features.

4.4 Knowledge-Based Recommendation

Overview: Knowledge-based systems use explicit rules or domain knowledge to filter or boost recommendations.
Implementation with H&M Dataset:

- **Features Used:** `product_type_name`, `colour_group_name`, `price`, `age`, `club_member_status`.

- **Method:**

  – Encode business rules (e.g., "Recommend jackets in winter", "Show member-only deals to ACTIVE members").

- Use price range to filter items for price-sensitive users.

- Apply color/style compatibility rules (e.g., "Do not recommend clashing colors in outfit suggestions").

- **Example:** If a user is under 25 and a club member, recommend "Divided" collection and exclusive offers.

Suitability:

- Ensures recommendations align with business goals and domain knowledge.

- Useful for cold start and special campaigns (e.g., holiday promotions).

- Increases trust and transparency for users.

4.5 Sequential and Context-Aware Recommendation (Optional Extension)

Overview: Sequential recommenders consider the order of user actions; context-aware systems use additional context (e.g., time, device).

Implementation with H&M Dataset:

- **Features Used:** `t_dat` (purchase date), device type (if available), session data (if available).

- **Method:**

  - Use session-based models (e.g., GRU4Rec) to predict next likely purchase.

  - Incorporate time-based features to recommend seasonal or trending items.

- **Example:** If a user bought "Swimwear" in June, recommend "Sunglasses" and "Sandals" in the same session.

Suitability:

- Captures short-term intent and seasonality.

- Enhances user experience with timely, relevant suggestions.

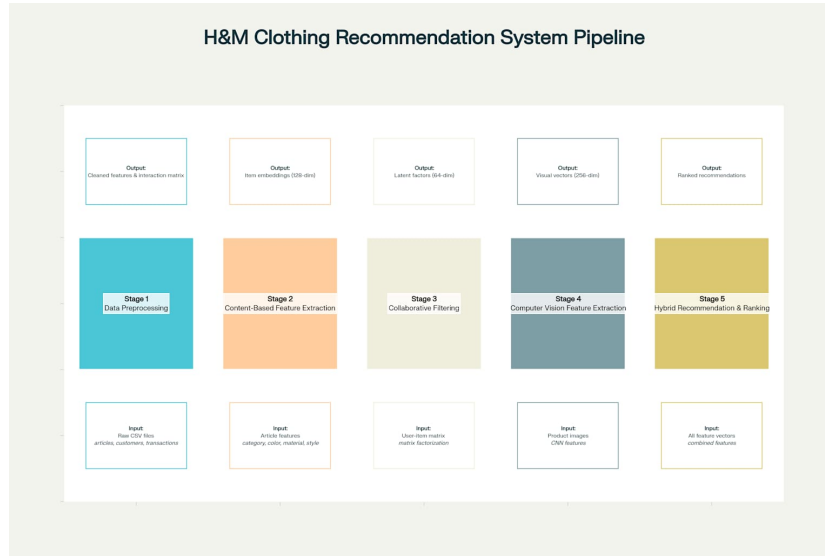## 4.6 Computer Vision-Based Recommendation (If Feasible)

Figure 4: Pipeline design and description for model with computer vision.

Overview: If all resources and time allow, computer vision can be integrated to extract visual features from product images, enabling recommendations based on visual similarity and style.

Implementation with H&M Dataset:

- **Features Used:** Product images (linked by `article_id`), plus all above metadata.

- **Method:**

  - Use a convolutional neural network (e.g., ResNet) to extract feature vectors from product images.

  - Compute visual similarity between items (e.g., using cosine similarity on image embeddings).

  - Combine visual similarity with content-based and collaborative scores in a hybrid model.

- **Example:** If a user likes a floral dress, recommend visually similar dresses (e.g., similar color, pattern, or silhouette) even if text metadata is sparse.

Suitability:

- Captures visual style and aesthetics, which are crucial in fashion.

- Addresses cold start for new items with images but little interaction data.

- Enables "search by image" and visually-driven recommendations.

4.7   Justification and Feasibility

- All proposed methods are supported by the H&M dataset's structure and fields.

- The combination of methods allows for robust evaluation and comparison.

- The plan is feasible within the project timeline, as each method can be implemented and tested incrementally.

- Hybrid and knowledge-based approaches ensure recommendations are both accurate and aligned with business needs.

- If time and resources allow, computer vision integration will further enhance recommendation quality, especially for visually-driven user preferences.

4.8 Summary Table: Methods and Dataset Field Usage

Table 11: Summary of Methods and Dataset Field Usage

Method	Key Fields Used	Why Suitable for H&M Dataset
Content-Based	product_type_name, colour_group_name, department_name, detail_desc	Leverages rich item metadata; good for new users/items; explainable
Collaborative Filtering	customer_id, article_id, t_dat	Captures user-item patterns; scales to large data; handles popularity
Hybrid	All above	Combines strengths; mitigates cold start; improves accuracy/diversity
Knowledge-Based	product_type_name, colour_group_name, price, age, club_member_status	Encodes business rules; supports campaigns; increases trust
Sequential/Context-Aware	t_dat, session/device (if available)	Models trends, seasonality, and session intent
Computer Vision	product images, article_id	Captures visual similarity and style; enables image-based search

In summary: A multi-method approach—combining content-based, collaborative, hybrid, knowledge-based, and (if feasible) computer vision recommenders—will best address the H&M recommendation problem, leveraging the dataset's strengths and supporting a wide range of user needs and business objectives.

5 Evaluation(s)

This section details the evaluation strategy for the H&M recommendation system, covering both model-level and system-level metrics, computational considerations, and the design of a user study. Concrete examples are provided to illustrate metric calculation and user feedback collection.

5.1 Model Evaluation Metrics (Offline, Historical Data)

To assess the accuracy and quality of the recommendation algorithms using historical purchase data, the following metrics will be used:

- **Precision@N:** The proportion of recommended items in the top-N list that are actually relevant (i.e., purchased by the user in the test set).

  - *Example:* If 3 out of 5 recommended items ($N = 5$) were actually purchased, then Precision@5 = 0.6.

- **Recall@N:** The proportion of relevant items that are recommended in the top-N list.

  - *Example:* If a user purchased 4 items and 2 appear in the top-5 recommendations, Recall@5 = 0.5.

- **NDCG@N (Normalized Discounted Cumulative Gain):** Measures ranking quality by giving higher weight to relevant items appearing higher in the recommendation list.

  - *Example:* If a relevant item is ranked 1st, it contributes more to NDCG than if ranked 5th.

- **Coverage:** The proportion of items/users for which the system can make recommendations.

- **Diversity:** Measures how varied the recommended items are (e.g., by category or color).

- **Novelty:** Assesses how often the system recommends items the user has not seen or purchased before.

Per-User Metrics: All metrics will be averaged over all users in the test set to ensure fairness and avoid bias toward highly active users.

5.2 System Evaluation Metrics (Online, User Interactions)

To evaluate the recommender system in a real or simulated environment, the following metrics will be tracked:

- **Click-Through Rate (CTR):** Fraction of recommended items that are clicked by users.

  - *Example:* If 100 recommendations are shown and 15 are clicked, CTR = 15%.

- **Conversion Rate:** Fraction of recommended items that are purchased.

- **Session Length:** Average number of interactions per user session.

- **Return Rate:** Fraction of recommended items that are returned by users (lower is better).

- **User Satisfaction:** Collected via post-session questionnaire (see below).

5.3 Metric Importance and Tradeoffs

- **Most Important:** Precision@N, Recall@N, and NDCG@N are prioritized for offline evaluation, as they directly measure recommendation relevance and ranking.

- **CTR and Conversion Rate** are most important for online/system evaluation, as they reflect real user engagement and business impact.

- **Tradeoffs:**

  - High precision may reduce diversity (recommending similar items).
  - High recall may lower novelty (recommending popular but already seen items).
  - The best model will be chosen based on a balanced improvement in Precision@N, NDCG@N, and online CTR, with minimum acceptable coverage and diversity.

5.4 Computational Requirements

- **Real-Time Recommendations:** The system must generate recommendations in under 200ms per user request to support interactive web/mobile interfaces.

- **Model Updates:** Models will be retrained weekly to incorporate new transactions and items. For cold start, content-based and knowledge-based methods will be used until sufficient data is available.

- **Scalability:** Efficient data structures (e.g., approximate nearest neighbor search for embeddings) and caching will be used to handle millions of users and items.

5.5 User Study Design

To evaluate the recommender system with real users, a simulated user interface will be developed. The study will involve the following steps:

1. **Participant Selection:** Recruit 10–20 users representing different segments (e.g., age, shopping frequency).

2. **Simulated Session:** Each user interacts with the system for 10–15 minutes, browsing and providing feedback on recommended items (e.g., like/dislike, add to cart).

3. **Data Collection:** Log all interactions (clicks, purchases, skips) for metric calculation.

4. **Post-Session Questionnaire:** Users complete a short survey to provide qualitative feedback.

Sample Questionnaire Items:

- How relevant were the recommendations to your style? (1–5)

- Did you discover any new items you liked? (Yes/No)

- How easy was it to use the recommendation interface? (1–5)

- Would you trust these recommendations for future purchases? (1–5)

- Any suggestions for improvement? (Open text)

5.6   Example: Metric Calculation

Suppose:

- User A is recommended 5 items: [A, B, C, D, E]

- User A actually purchased: [B, D, F]

Then:

- Precision@5 = 2/5 = 0.4 (B and D are relevant)

- Recall@5 = 2/3 0.67

- If B is ranked 1st and D is ranked 4th, NDCG@5 will be higher than if both are ranked lower.

5.7 Summary Table: Metrics Overview

Table 12: Evaluation Metrics for H&M Recommendation System

Metric	Type	Purpose/Example
Precision@N, Recall@N, NDCG@N	Model (offline)	Relevance and ranking of recommendations; e.g., Precision@5 = 0.4
Coverage, Diversity, Novelty	Model (offline)	System robustness and user experience; e.g., 90% coverage, 0.7 diversity score
CTR, Conversion Rate, Session Length	System (online)	User engagement and business impact; e.g., CTR = 15%
Return Rate	System (online)	Business cost and satisfaction; e.g., 5% of recommended items returned
User Satisfaction	System (user study)	Qualitative feedback; e.g., average rating 4.2/5

6 Conclusion

This project design outlines a thorough and effective strategy for developing a strong fashion recommendation system for H&M, utilising real-world datasets and a range of recommendation techniques suited to different user preferences. The proposed system effectively combines various techniques, including content-based, collaborative, hybrid, knowledge-based, and potentially computer vision, to provide precise, varied, and captivating recommendations via scalable web and mobile interfaces. The evaluation plan integrates

robust offline metrics with user-focused online assessments and a simulated user study, guaranteeing that both technical performance and practical usability are thoroughly considered. This framework offers a straightforward guide for execution, continuous enhancement, and significant influence on user satisfaction and organisational results.