UNIT - 1

Recommender System functions:

Recommender systems are algorithms designed to suggest relevant items to users based on various types of data.

Here are some key functions and components of recommender systems:

1. Data Collection:
   * User Data:
     Information about the user such as demographics, behaviour, preferences etc.,
   * Item Data:
     Information about items such as characteristics, descriptions categories etc.,
   * Interaction Data:
     Records of Interactions between users and items such as ratings, clicks, purchases etc.,

2. Data Processing:

* Normalinzation:
Adjusting data to a standard scale (e.g., rating normalinzation).
* Feature Extraction:
Identifying and extracting relevant features from raw data.
* Dimensionality Reduction:
Reducing the number of features to simplify the model. (E.g., using PCA).

3. Modeling:

* Collaborative Filtering: Recommending items based on the behaviour of similar users or items.
    * User-based Collaborative Filtering:
    Finding users similar to the target user and recommending items they liked.
    * Item-based Collaborative filtering:
    Finding items similar to those the target user liked and recommending these items

Jan

Appointment Calls to Reply Important Calls

* Content-based Filtering:
Recommending items similar to those the user has shown interest in, based on item features.
* Hybrid Methods:
Combining collaborative and content-based methods to leverage the strengths of both.

4. Training and Optimization:
* Model Training:
Using historical data to train machine learning models to predict user preferences.
* Hyperparameter Tuning:
Optimizing parameters of the model to improve performance.
* Cross-validation:
Assessing the model's performance on different subsets of data to ensure it generalizes well.

5. Recommendation Generation:
* Top-N Recommendations:
Generating a list of the top N items for a user.

* Ranking Algorithms:
Ranking items based on predicted user preferences.
* Contextual Recommendations:
Considering context such as time, location, or device to generate recommendations.

6. Evaluation and Metrics:
* Precision and Recall:
Measuring the accuracy of the recommendations.
* F1 Score:
A balance between precision and recall.
* Mean Absolute Error (MAE):
Measuring the average magnitude of errors in predictions.
* Root Mean Squared Error (RMSE):
Measuring the square root of the average squared differences between predicted and actual ratings.
* Diversity and Novelty:
Assessing how varied and new the recommendations are to the user.

7. Deployment and Monitoring:

* Real-time Recommendations:
  Generating recommendations on-the-fly as the user interacts with the system.
* Batch Processing:
  Periodically updating recommendations based on the latest data.
* A/B Testing:
  Comparing different recommendation strategies to determine the most effective one.
* Feedback Loop:
  Incorporating user feedback to continuously improve the recommendation model.

8. Personalization:

* User Profiling:
  Building profiles for users based on their interactions and preferences.
* Contextual Personalization:
  Tailoring recommendations based on the current context of the user.

9. Scalability:

* Efficient Algorithms:
  Ensuring algorithms can handle large amounts of data and many users/items.
* Distributed Systems:
  Using distributed computing to process data and generate recommendations at scale.

10. Security and Privacy:

* Data Anonymization:
  Protecting user identities while using their data for recommendations.
* Secure Data Storage:
  Ensuring user data is stored securely and access is controlled.

By integrating these functions, recommender systems can provide personalized, relevant, and engaging experiences for users across various domains such as e-commerce, entertainment, social media and many more.

Linear Algebra Notation: Matrix Addition:

Jan

Linear Algebra:
A branch of mathematics concerning linear equations and their representations using vectors and matrices.

Matrices:
Rectangular arrays of numbers, symbols, or expressions arranged in rows and columns.

Importance:
Fundamental in various scientific fields including physics, computer science, engineering, and economics.

Matrix Addition:

Definition:
Sum of two matrices of the same dimension by adding their corresponding entries.

Matrix $A = \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 4 \\ 5 & 6 \end{bmatrix}$

$$A + B = \begin{bmatrix} 2+1 & 3+4 \\ 7+5 & 8+6 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 12 & 14 \end{bmatrix}$$

Appointment

Calls to Reply

Important Calls

Properties:

Commutative: $A+B = B+A$

Associative : $(A+B)+C = A+(B+C)$

Matrix Multiplication:

Definition:

Product of two matrices, where the number of columns in the first matrix equals the number of rows in the second matrix.

Notation : $A \times B$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 10 & 11 \\ 20 & 21 \\ 30 & 31 \end{bmatrix}$$

$$A \times B = \begin{bmatrix} 1\times 10 + 2\times 20 + 3\times 30 & 1\times 11 + 2\times 21 + 3\times 31 \\ 4\times 10 + 5\times 20 + 6\times 30 & 4\times 11 + 5\times 21 + 6\times 31 \end{bmatrix}$$

$$= \begin{bmatrix} 10+40+90 & 11+42+93 \\ 40+100+180 & 44+105+186 \end{bmatrix} = \begin{bmatrix} 140 & 146 \\ 320 & 335 \end{bmatrix}$$

Not Commutative:

$$AB \neq BA$$

Associative:

$$(A \times B) \times C = A \times (B \times C)$$

Distributive:

$$A \times (B + C) = A \times B + A \times C$$

Matrix Transposition:

Matrix transposition is a fundamental linear algebra operation that flips a matrix by interchanging its rows and columns. This transformation can be useful in various applications, such as data analysis, image processing and machine learning.

Notation:

The transpose of a matrix A is denoted as A^T.

Superscript T indicates that the rows and columns have been swapped.

Visualization:

Transposing a matrix can be visualized by flipping the matrix along its main diagonal turning rows

into columns and vice versa.

Properties of Matrix Transposition.

Symmetry;
The transposed matrix is symmetric to the original matrix across the main diagonal, meaning the elements are flipped across this axis.

Dimensions:
m rows will have m columns
n columns will have n rows.

Inverse Relationship:
Taking the transpose of a matrix twice returns the original matrix. (inverse nature of this operation)

Linearity:
Transposition is a linear operation, meaning it distributes across matrix addition and scalar multiplication.

Applications of Matrix Transposition:

* Data Analysis:
  - helps organize and manipulate large datasets more effectively.

Image Processing:

transposing matrices is used to rotate, flip and transform digital images in various ways.

Linear Algebra:

To solve systems of linear equations and computing determinants.

Transposition and Matrix operations:

* Commutative Property:

$$(A+B)^T = A^T + B^T$$

* Distributive Property:

$$(AB)^T = B^T \cdot A^T$$

* Identity Matrix:

The transpose of the identity matrix I is the identity matrix itself i.e., $I^T = I$.

* Inverse Matrices:

If A is an invertible matrix, then $(A^{-1})^T = (A^T)^{-1}$ allowing us to efficiently compute inverse matrices.

Matrix Inverse:

The inverse of matrix is the matrix that on multiplying with the original matrix results in an identity matrix. For any matrix A, its inverse is denoted as A^{-1}.

$$A^{-1} = \frac{1}{|A|} \text{Adjoint } A$$

where

$|A|$ = determinant

For matrix A and its inverse of A^{-1}, the identity property holds.

$$A \cdot A^{-1} = A^{-1} A = I$$

where

I is the identity matrix.

Singular Matrix:

A matrix whose value of the determinant is zero is called a singular matrix. i.e., any matrix A is called singular matrix if $|A| = 0$. Inverse of a singular matrix does not exist.

Non-Singular Matrix:

A matrix whose value of the determinant is non-zero is called a non-singular matrix, i.e., any matrix A is called a non-singular matrix if $|A| \neq 0$. Inverse of a non-singular matrix exists.

Identify Matrix:

A square matrix in which all the elements are zero except for the principal diagonal elements is called the identify matrix. It is represented using I. It is the identify element of the matrix as for any matrix A.

$$A \times I = A$$

$$I_{3\times3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Inverse of a Matrix Formula:

The inverse of Matrix A, that is A^{-1} is calculated using the inverse of matrix formula which involves dividing the adjoint of a matrix by its determinant.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Inverse of Matrix Formula:

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$A^{-1} = \frac{\text{Adj } A}{|A|}$$

(Eg.) Find the inverse of the following matrix:

$$A = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix} \qquad A^{-1} = \frac{1}{\det(A)}$$

Jan

$$A^{-1} = \frac{1}{8-9}$$

$$A^{-1} = \frac{1}{-1}\begin{bmatrix} 2 & -3 \\ -3 & 4 \end{bmatrix}$$

$$A = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix} \qquad A^{-1} = \frac{1}{-1}\begin{bmatrix} 2 & -3 \\ -3 & 4 \end{bmatrix}$$

$$A A^{-1} = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix}$$

$$= \begin{bmatrix} -8+9 & \cancel{6+9}\ 12-12 \\ -6+6 & 9-8 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$ → This identity matrix confirms inverse relationship.

Covariance Matrix:

Covariance Matrix is a type of matrix used to describe the covariance values between two items in a random vector. It is also known as variance-covariance matrix because the variance of each element is represented along the matrix's major diagonal and the covariance is represented among the non-diagonal elements.

The co-variance matrix is a square matrix whose variable can take any value - positive, negative or zero.

A positive covariance suggests that the two variables have a positive relationship.

A Negative covariance indicates that they do not have a positive relationship.

Zero covariance means that there is no relationship.

(Eg.) Let's say there are 2 data sets $X = [10, 5]$ and $Y = [3, 9]$

The variance of set $X = 12.5$

set $Y = 18$

Appointment Calls to Reply Important Calls

The covariance between both variables is -15.

The covariance matrix is:

$$= \begin{bmatrix} 12.5 & -15 \\ -15 & 18 \end{bmatrix}$$

Feb

How to find Covariance Matrix?
Consider the following data:

Student	Maths (x)	Physics (y)
Ganesh	80	70
Mahesh	63	20
Suresh	100	50

The following steps have to be followed:

Step 1:
Find the mean of variable x.

Appointment | Calls to Reply | Important Calls

Sum up all the observations in variable x and divide the sum obtained with the number of terms. Thus

$$\frac{(80+63+100)}{3} = 81$$

Step 2:

Substract the mean from all observations.

$(80-81)$, $(63-81)$, $(100-81)$.

Step- 3:

Take the squares of the differences obtained above and then add them up.

Thus,

$$(80-81)^2+(63-81)^2+(100-81)^2$$

Step-4:

Find the variance of x by dividing the value obtained in step 3 by 1 less than the total no. of observations.

$$var(x) = \frac{\left[(80-81)^2+(63-81)^2+(100-81)^2\right]}{(3-1)}$$

$$= 343$$

Step 5:

Similarly, repeat steps 1 to 4 to calculate the variance of Y.

$$Var(Y) = 633.333$$

Step 6:

choose a pair of variables.

Step 7:

subtract the mean of the first variable (X) from all observations.

$$(80-81), (63-81), (100-81).$$

Step 8:

Repeat the same for variable Y; $(70-47)$ $(20-47)$, $(50-47)$.

Step 9:

Multiply the corresponding terms:

$$(80-81)(70-47), (63-81)(20-47), (100-81)(50-47)$$

Step 10:

Find the covariance by adding these values and dividing them by $(n-1)$.

$$COV(X,Y) = \frac{\left[(80-81)(70-47)+(63-81)(20-47)+(100-81)(50-47)\right]}{(3-1)}$$

$$= 260.$$

Step II:

Use the general formula for the covariance matrix to arrange the terms. The matrix becomes:

$$\begin{bmatrix} 343 & 260 \\ 260 & 633.333 \end{bmatrix}$$

Properties of Covariance Matrix:

* A covariance matrix is always square, implying that the number of rows in a covariance matrix is always equal to the number of columns in it.

* A covariance matrix is always symmetric, implying that the transpose of a covariance matrix is always equal to the original matrix.

* A covariance matrix is always positive and semi-definite.

Appointment | Calls to Reply | Important Calls

* The eigenvalues of a covariance matrix are always real and non-negative.

Applications of Recommendation Systems :

Recommendation systems are widely used across various domains to enhance user experience, drive engagement, and increase sales. Here are some key applications:

1. E-commerce

- **Product Recommendations**: Suggesting products to users based on their browsing history, purchase history, and preferences. Examples include Amazon's "Customers who bought this also bought" feature.
- **Personalized Discounts and Offers**: Providing customized discounts and promotional offers based on user behavior and preferences.

2. Entertainment

- **Movie and TV Show Recommendations**: Platforms like Netflix and Hulu use recommendation systems to suggest movies and TV shows based on users' viewing history and ratings.
- **Music Recommendations**: Services like Spotify and Apple Music recommend songs, albums, and artists based on listening history and user preferences.

3. Social Media

- **Content Recommendations**: Social media platforms like Facebook, Instagram, and Twitter use recommendation systems to show personalized content, such as posts, articles, and videos, in users' feeds.
- **Friend Suggestions**: Algorithms suggest new friends or connections based on mutual friends, interests, and interactions.

4. News and Information

- **Article Recommendations**: News websites and apps recommend articles based on users' reading history and interests. Examples include Google News and Flipboard.
- **Personalized News Feeds**: Aggregating and prioritizing news stories that are relevant to individual users.

5. Online Advertising

- **Targeted Advertising**: Serving personalized ads based on users' browsing history, search queries, and demographic information. Platforms like Google Ads and Facebook Ads use sophisticated recommendation algorithms for ad targeting.

6. Education

- **Course Recommendations**: Online learning platforms like Coursera and Udemy recommend courses based on users' learning history, interests, and career goals.
- **Learning Path Recommendations**: Suggesting personalized learning paths and resources to help users achieve specific learning objectives.

7. Healthcare

- **Medical Recommendations**: Recommending treatments, health tips, and preventive measures based on patients' medical history and health data.
- **Personalized Health Plans**: Tailoring health and wellness plans based on individual health profiles and preferences.

8. Travel and Hospitality

- **Destination and Activity Recommendations**: Suggesting travel destinations, hotels, and activities based on users' travel history and preferences. Examples include TripAdvisor and Airbnb.
- **Itinerary Planning**: Creating personalized travel itineraries based on user interests and preferences.

9. Finance

- **Investment Recommendations**: Offering personalized investment advice and product recommendations based on users' financial goals, risk tolerance, and portfolio performance.
- **Credit and Loan Offers**: Recommending credit cards, loans, and other financial products that match users' credit profiles and needs.

10. Retail Banking

- **Personalized Banking Services**: Tailoring banking services and product recommendations based on individual customer behavior and financial status.

11. Gaming

- **Game Recommendations**: Suggesting games to users based on their playing history and preferences. Examples include Steam and Xbox Live.
- **In-game Content Recommendations**: Recommending in-game purchases, levels, and items based on user behavior.

12. Job Portals

- **Job Recommendations**: Suggesting job listings to job seekers based on their resume, search history, and application behavior. Examples include LinkedIn and Indeed.
- **Candidate Recommendations**: Helping recruiters find potential candidates by recommending profiles that match job requirements.

13. Real Estate

- **Property Recommendations**: Suggesting properties to users based on their search criteria and browsing history. Examples include Zillow and Redfin.
- **Personalized Property Alerts**: Notifying users about new listings that match their preferences.

14. Restaurants and Food Delivery

- **Restaurant Recommendations**: Recommending restaurants based on users' location, dining history, and preferences. Examples include Yelp and Zomato.

- **Dish Recommendations**: Suggesting specific dishes or meal plans based on users' dietary preferences and order history.

Recommendation systems leverage data and machine learning to create personalized experiences across these diverse applications, ultimately driving user satisfaction and engagement.

Issues in Recommender Systems :

Recommender systems, despite their widespread use and benefits, face several challenges and issues. Here are some of the key issues:

1. Data Sparsity

- **Limited User Interactions**: Many users interact with only a small subset of items, leading to sparse data matrices. This makes it difficult to accurately predict user preferences.
- **Cold Start Problem**: New users and new items lack interaction data, making it challenging to generate accurate recommendations for them.

2. Scalability

- **Large Datasets**: As the number of users and items grows, the computational requirements for generating recommendations increase significantly.
- **Real-Time Recommendations**: Generating recommendations in real-time for a large number of users can be computationally expensive and complex.

3. Diversity and Novelty

- **Lack of Diversity**: Recommender systems often suggest items similar to those a user has already interacted with, leading to a narrow range of recommendations and reducing the serendipity of discovering new items.
- **Balancing Novelty**: Introducing novel items while still maintaining relevance is a challenging task. Too much novelty can reduce user satisfaction if the recommendations are not closely aligned with their preferences.

4. Bias and Fairness

- **Algorithmic Bias**: Recommender systems can inadvertently reinforce existing biases present in the data, leading to unfair treatment of certain user groups or item categories.
- **Popularity Bias**: Frequently recommended items tend to be popular ones, which can overshadow less popular but potentially relevant items.

5. Privacy Concerns

- **Data Collection**: Recommender systems often require extensive data collection, raising concerns about user privacy and data security.
- **User Trust**: Ensuring that users trust the system to handle their data responsibly is crucial for the long-term success of recommender systems.

6. Evaluation and Metrics

- **Evaluation Metrics**: Selecting appropriate metrics to evaluate the performance of recommender systems can be challenging. Common metrics like accuracy may not fully capture the quality of recommendations.
- **A/B Testing**: Conducting reliable A/B tests to measure the impact of recommendation algorithms in real-world settings can be complex and resource-intensive.

7. Interpretability and Transparency

- **Black Box Models**: Many recommendation algorithms, especially those based on deep learning, are difficult to interpret, making it hard to understand why certain recommendations are made.
- **User Understanding**: Providing explanations for recommendations can improve user trust and satisfaction, but generating meaningful and understandable explanations is challenging.

8. Dynamic Preferences

- **Changing Interests**: Users' preferences can change over time, and recommender systems need to adapt quickly to these changes to remain relevant.
- **Temporal Dynamics**: Capturing and modeling the temporal aspects of user behavior and preferences is complex.

9. Content Limitations

- **Content Quality**: The quality and completeness of item descriptions and metadata can significantly impact the performance of content-based recommendation systems.
- **Multimedia Content**: Recommending items such as images, videos, and music requires sophisticated content analysis techniques, which can be computationally intensive.

10. Context-Awareness

- **Contextual Factors**: Incorporating contextual information (e.g., location, time of day, device used) into recommendations can enhance relevance but adds complexity to the modeling process.
- **Personal Context**: Understanding and leveraging personal contextual factors such as mood or situational context can further improve recommendations.

Addressing these issues requires a combination of advanced algorithms, robust data management practices, and a careful consideration of ethical and user-centric design principles.