Emotion-Aware Movie Recommendation System Using Neural Networks and NLP

Project report submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Technology

by

Bhavya Ameta - 22UCS047(CSE)
Anushka Singh - 22UCS029(CSE)
Bhavik Mehta - 22UCS046(CSE)

Under Guidance of
Rajendra Shivaji Patil (Department of CSE)

LNMIIT
The LNM Institute of
Information Technology

The LNM Institute of Information Technology, Jaipur

April 2025

The LNM Institute of Information Technology
Jaipur, India

CERTIFICATE

This is to certify that the project entitled "Emotion-Aware Movie Recommendation System Using Neural Networks and NLP" , submitted by Bhavya Ameta (22UCS047), Anushka Singh(22UCS029) and Bhavik Mehta(22UCS046) in partial fulfillment of the requirement of degree in Bachelor of Technology (B. Tech), is a bonafide record of work carried out by them at the Department of Computer Science and Engineering, The LNM Institute of Information Technology, Jaipur, (Rajasthan) India, during the academic session 2024-2025 under my supervision and guidance and the same has not been submitted elsewhere for award of any other degree. In my/our opinion, this thesis is of standard required for the award of the degree of Bachelor of Technology (B. Tech).

Date

Adviser: Name of BTP Supervisor

Acknowledgments

We sincerely thank Dr. Rajendra Patil for guiding us throughout our project span. His contributions were invaluable in shaping the direction and success of our work. We are deeply grateful for his insightful feedback and unwavering support.

Abstract

The ”Emotion-Aware Movie Recommendation System” project seamlessly blends Neural Network and Natural Language Processing to transform how movie recommendations are handled for film viewers.

Traditional recommendation systems have, until now, used simple features like content already watched, content rated, or simple genre classification. However, these traditional methods overlook a very fundamental aspect of the human experience: the evolving nature of our feelings. The suggested system bridges this very basic gap by putting feelings at the heart of the recommendation process. As our users engage with our site, they give us natural language inputs that are put through analysis by our top-of-the-line BERT-based model. This model, which has been thoroughly fine-tuned to perfection on the high-quality GoEmotions dataset, detects and quantifies the user’s emotional state at that instant along multiple dimensions. Instead of translating these emotions into data points, our system adeptly projects them onto resonating and contrasting genres of movies and thereby creates a mapping of emotions genres of the most likely-to-engage content at that instant.The system also asks the user whether they wish to shift their current mood. If so, it adjusts the genre weighting accordingly by increasing the weight of either resonant or contrastive genres in the recommendation model.

The recommendation model is trained on the vast MovieLens 32M dataset, which has a great deal of information related to many movies. Our method is grounded in a Neural Collaborative Filtering model that considers three main inputs: userID, movieID, and an affective dimension of novelty genre preference. Through the inclusion of these factors with advanced embedding methods and multi-hot encoding for genre data, we generate highly accurate recommendations considering not only past tastes but also the emotional state of the user at the moment.

The emotion-based recommendation is a major development in recommendation technology. Through recognition of the strong impact of emotional states on viewing actions, we make recommendations that feel naturally suited to consumers’ tastes and accuracy. The combination of deep learning methods improves the accuracy and relevance of the suggestions, so that recommendations actually correlate with users’ current emotional states. This present work is an important contribution to the new area of emotionally intelligent artificial intelligence systems, demonstrating that sentiment analysis and collaborative filtering algorithms can be merged to generate more human-sensitive technology. In an era when our online activities seem to grow more impersonal by the minute, our system restores the useful human intuition that emotions have an important role in determining our likes and dislikes and thus offers a more authentic and rewarding way to discover movies we will like in real time.

Contents

Chapter Page

1 . . . 1
  1.1 Introduction to the Field . . . 1
  1.2 The Area of Work . . . 2
  1.3 Motivation . . . 2
  1.4 Problem Addressed . . . 4
  1.5 Existing System . . . 4
    1.5.1 Netflix . . . 4
    1.5.2 MoodMovie . . . 5

2 Literature Survey . . . 6
  2.1 Introduction . . . 6
  2.2 Traditional Recommendation Approaches . . . 6
    2.2.1 Collaborative Filtering . . . 6
    2.2.2 Content-Based Filtering . . . 7
    2.2.3 Matrix Factorization and Latent Factor Models . . . 7
  2.3 Deep Learning Approaches for Recommendation . . . 8
    2.3.1 Neural Collaborative Filtering (NCF) . . . 8
    2.3.2 Sequence-Aware Recommendation Models . . . 10
    2.3.3 Autoencoder-Based Approaches . . . 11
    2.3.4 Emotion-Aware Recommendation Systems . . . 11
  2.4 Our Current Implementation . . . 12
    2.4.1 Hybrid Nature of Our Current Model . . . 12

3 Proposed Work . . . 14
  3.1 Content-Based vs. Collaborative Filtering . . . 14
  3.2 Alternatives Explored and Why They Were Dropped . . . 14
    3.2.1 Autoencoders . . . 14
    3.2.2 Transformers (e.g., BERT4Rec) . . . 14
    3.2.3 Matrix Factorization . . . 15
  3.3 **Final Architecture** . . . 15
  3.4 Technical Background . . . 16

4 Simulation and Results . . . 19
  4.1 Results . . . 19
  4.2 Simulation . . . 20

5 22
5.1 Conclusions 22
5.2 Scope of further work 22
5.2.1 **What We Will Do** 22
5.2.2 **What Can Be Done** 22

Bibliography 24

Chapter *1*

1.1 Introduction to the Field

In the current digital era, individuals are bombarded with an excessive amount of choices—anything from choosing a movie to watch, a song to hear, a book to read, or a product to purchase. The excess of choices has created the need for recommendation systems. These systems simplify the decision-making process by providing personalized recommendations based on user behavior, preferences, and historical patterns. Netflix, Spotify, and Amazon, among others, depend on recommendation algorithms to increase user engagement by personalizing content to satisfy individual tastes.

In the movie recommendation space, two techniques are widely used:

- **Collaborative Filtering** makes use of similar users' exhibited tastes. For example, if User A and User B both exhibited similar tastes in the past, a film liked by User A but not by User B so far can be suggested to User B. The method operates on the principle that "people with similar tastes have liked this too."

- **Content Based Filtering** is based on focusing on product features and individual users' preferences, however. In case a user continuously watches romantic comedies starring a leading lady, the system continuously suggests similar films regardless of the behavior of other users.

  Though proven to work empirically, such techniques merely have little regard for the dominant human factor: **emotionality**. Using media is not always a case of reason or habit but is actually intrinsically linked to emotions. One might choose an inspirational film on a stressful working day or a sad drama apt for a bad mood .Emotional needs—either mood congruence (enforcement of the existing mood) or mood management (shift to a desired emotional state)—play an important role in content selection. Traditional systems usually do not have this emotional awareness. With the aim to overcome this lack, we introduce an emotion-aware movie recommendation system that captures emotions from user-input text (e.g., journaling, reviews, or mood descriptions) and tailors movie recommendations based on these emotions.

1.2 The Area of Work

Based on the shortfalls that exist in classical recommender systems, this project proposes an emotion-sensitive movie recommender system based on the convergence of natural language processing (NLP) and neural collaborative filtering (NCF) to provide very personalized and emotionally sensitive suggestions.

At the core of our system lies a Transformer-trained model on the GoEmotions dataset, which allows it to learn subtle expressions of emotions within user content. As users are on the platform, they are asked to state their emotional intent—whether they want to carry on with the current mood or switch to another. This simple yet effective choice has a direct influence on the way we rank genres in our recommendation system:

- If the user wants to alter their emotional state, we give contrasting genres a weight of 0.7, while resonatng genres that match the prevailing emotion get a weight of 0.3.
- If the user wants to keep the rating level, the weighting is reversed—resonant genres are given the 0.7 weight, and contrast genres are given 0.3.
- Where users remain neutral in terms of emotions, we offer an equal weight of 0.5 for both resonant and contrasting genres.

Furthermore, we also suggest another subtle but significant improvement: when a genre appears in both the resonant and contrast mappings, we average their weights to increase their combined effect on the final recommendation score.

These affectively charged genre choices, user profiles, and metadata from movies are all embedded and analyzed by our deep learning-based NCF model, which is trained on the vast MovieLens 32M dataset. This blend of emotional awareness and collaborative filtering is a departure towards user-oriented design—one that recommends not just based on historical behavior but also on the emotional experience of the user in the present. Our work sits at the intersection of affective computing, recommendation systems, and human-centered design, producing recommendations that are not only relevant but also affectively informed—enabling comfort, resonance, or change depending on the state and intent of the user.

1.3 Motivation

The inspiration for this work derives from an abundance of observations:

- **Emotion-based content consumption:** Research in media psychology indicates that emotions play a controlling role in the choice of content, yet most recommendation systems only take into account past behavior and explicit preferences.

- **Context-aware suggestions:**Traditional systems provide content recommendations that are not sensitive to the immediate emotional context of the user and therefore provide irrelevant suggestions.

- **Psychological well-being:** Providing material that satisfies emotional needs can directly enhance user satisfaction and psychological well-being by emotional bonding or emotion regulation.

- **Building better personalization:** By adding contextual emotional content, recommendation systems are able to present more sophisticated personalization than demographics or preference profiles.

- **Cross-discipline innovation:** This project combines natural language processing, emotion detection, machine learning, and media psychology, thus creating a pioneering cross-disciplinary approach in recommendation-making.

In the age of contemporary digital entertainment, recommendation systems have emerged as a key component making the highest contribution towards user experience and satisfaction. Netflix, YouTube, and Amazon owe a significant part of their success to their recommendation algorithms—that drive 80% of Netflix viewership, 60% of the front page traffic on YouTube, and 35% of Amazon sales. Yet, although incredibly effective, these recommendation systems are built mainly on static past information (e.g., past view counts and ratings) and group behavior patterns, and completely ignore two vitally inherent inputs:

- **Real-time emotional context –** A user's current mood significantly influences their content preferences.

- **Group dynamics –** Multi-user scenarios (e.g., groups of friends/family watching together) introduce difficult preference conflicts that can't be addressed by traditional systems.(future work)

To fill such gaps, we **introduce an emotion-aware, NCF-based recommendation system based on NLP for real-time emotion detection**. Compared to traditional methods, our system specifically incorporates:

- Text-based emotion detection (with a GoEmotions-trained bert model) to predict user mood from input sentences (e.g., "I'm feeling stressed" $\rightarrow$ "sad/anxious").

- Emotion-genre mapping (e.g., "happy" $\rightarrow$ Comedy, "sad" $\rightarrow$ Drama) in order to adjust suggestions adaptively.

Neural Collaborative Filtering (NCF) trained on the MovieLens 32M dataset to improve emotion-specified genre recommendations. Aggregate group sentiment to resolve multi-user conflicts (future work).

1.4 Problem Addressed

The problem this project addresses can be stated as follows:

How do a movie recommendation system properly use users' existing emotional states, inferred from text input, to make contextually and collaboratively relevant movie suggestions that resonate with or change users' emotions?

Specific challenges to be solved include:

- Accurately extracting emotions from text input by users
- Combining emotion detection with collaborative filtering for personalized recommendations
- Striking a balance between previous tastes and current emotional state
- Providing users with the ability to select whether the recommendations should agree with or switch their current mood
- Developing a system that functions effectively under varying states of emotions and user preferences

1.5 Existing System

Several movie recommendation systems are deployed, which are mainly based on collaborative filtering, content-based filtering, or their hybrids. Though these systems perform well in determining user preferences through past behavior and ratings, they lack in interpreting **the emotional context** of the user while interacting. Incorporating a user's **present emotional state**, particularly from natural language input, is largely an unexplored area in conventional systems.

1.5.1 Netflix

Netflix's recommendation engine is central to content discovery by providing extremely personalized recommendations to viewers. The system relies on a mix of user behavior analysis, content metadata, and machine learning to suggest movies and TV shows.

Working of Netflix

Netflix's system uses a **hybrid model** that includes:

- **Collaborative Filtering**:Suggests content based on user similarities and what other similar users like.
- **Content-Based Filtering**: Suggests content that is comparable to the content a user has consumed in the past, based on metadata like genres, actors, directors, and themes.

- **Contextual Bandits**: These are utilized to strike a balance between exploration of new content and exploitation of known tastes, depending on time of day, device, or other contextual situations.
- **Deep Learning Models**: Neural networks (such as CNNs, RNNs, and VAEs) are utilized to analyze user activity patterns, metadata of text, and even thumbnail images for higher engagement.
- **Personalized Artwork and Home Page Layout**: Various thumbnails and row configurations are displayed to various users based on their interactive history.

  The recommendation system of Netflix is not able to include real-time emotional user feedback, and thus it cannot be as adaptable to the current mood of users. It doesn't provide a facility for users to indicate if they prefer to watch content that aligns with or changes their emotional state and instead uses historical viewing behavior, which might not always be compatible with the current emotional needs of the user.

1.5.2 MoodMovie

MoodMovie is a research prototype that tries to recommend movies according to user moods. Users manually pick among a set of emotional states, and the system suggests movies associated with every mood.

Working of MoodMovie

- User selects a mood tag from predefined options (e.g., sad, excited, romantic).
- Every mood is associated with a range of corresponding genres or themes.
- The system retrieves movies that correspond to the genre-mood mapping.

The system's emotion identification is done manually and is not real-time automated. It has no support for free-text emotion analysis and can't query users if they would like to preserve or change their mood. Furthermore, the fixed emotion-genre mapping constrains personalization and flexibility and isn't very sensitive to dynamic emotional states .

Chapter *2*

Literature Survey

2.1 Introduction

Recommendation systems have grown popularity, helping users discover content aligned with their preferences. In the domain of movie recommendations, traditional approaches have primarily focused on user-item interactions and content similarities. However, these methods often ignore the emotional context that controls user preferences at different times. This literature review looks into the evolution of recommendation techniques that channeled our approach to developing an emotion-aware movie recommendation system that integrates neural collaborative filtering with emotion analysis.

2.2 Traditional Recommendation Approaches

2.2.1 Collaborative Filtering

Collaborative filtering (CF) is one of the fundamentals techniques used in recommendation systems. It is based on the hypothesis that past agree users on evaluating the items would continue to agree [15].

Two primary variants of collaborative filtering were explored:

- **User-Based Collaborative Filtering:** It suggests selecting those users having similar rating patterns to a target user and suggests items which the similar users have rated high. Similarity among users is usually computed in a measure such as Pearson correlation or cosine similarity [5].

- **Item-Based Collaborative Filtering:** This method calculates similarities among items according to users' ratings and suggests items similar to those the user has rated positively before [14]. This approach is particularly valuable for domains with more stable item relationships than user preferences.

The main limitation of these traditional CF approaches is their inability to address the cold-start problem efficiently and to capture complex nonlinear relationships between users and items.

2.2.2 Content-Based Filtering

Content-based filtering techniques create recommendations through matching of user preferences with item features [10]. In movie recommendation, these features may be movie genres, actors, directors, and other metadata.

Major techniques explored under content-based filtering were:

1. **Bag-of-Words (BoW):**This technique models textual content (e.g., movie descriptions) as numerical feature vectors, which can be used to compute similarity using vector operations. Every dimension is one of the terms in the vocabulary, and the value is the importance of the term .

2. **TF-IDF (Term Frequency-Inverse Document Frequency):** This quantitative measure calculates the importance of a word to a document in a collection by rising in proportion to its frequency within the document but discounted by the frequency of the word in the corpus [13].

3. **Cosine Similarity:** This measure calculates the cosine of the angle between two vectors in an n- dimensional space with a similarity score ranging from -1 to 1. It is especially well-suited for comparing document vectors in content-based recommendations [17].

Content-based approaches are best at suggesting items that are like those a user has liked before but tend to suffer from overspecialization, constraining the diversity of suggestions.

2.2.3 Matrix Factorization and Latent Factor Models

Matrix Factorization (MF) techniques represent a significant advancement in recommendation systems. These approaches decompose the user-item interaction matrix into lower-dimensional matrices representing latent factors [8]. The factorization process captures implicit relationships between users and items beyond explicit features.

Key aspects of Matrix Factorization explored include:

1. Singular Value Decomposition (SVD): A mathematical technique that decomposes a matrix into three components, allowing dimensionality reduction while preserving the most significant information in the original matrix [3].

2.Probabilistic Matrix Factorization (PMF): A probabilistic extension of SVD that assumes Gaussian observation noise and defines a generative model for the ratings [11]

3. Alternating Least Squares (ALS): An algorithm that alternates between fixing the user factors and item factors, solving a more manageable least squares problem at each step [22].

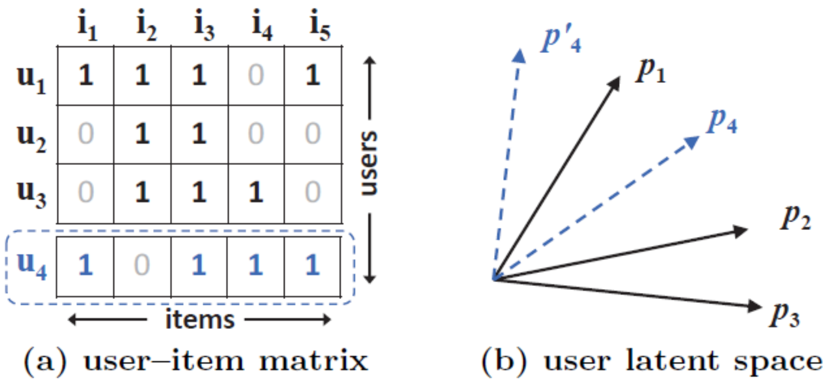


Figure 1: An example illustrates the limitation of MF. From data matrix (a), u4 is most similar to u1, followed by u3, and lastly u2. However, in the latent space (b), placing p4 closest to p1 makes p4 closer to p2 than p3, incurring a large ranking loss[4].

Matrix factorization approaches demonstrated superior performance compared to traditional collaborative filtering methods, particularly for sparse datasets. However, they still face challenges in capturing complex non-linear relationships and integrating additional contextual information.

2.3 Deep Learning Approaches for Recommendation

The emergence of deep learning has transformed recommendation systems by enabling more complex modeling of user-item interactions. Several neural network architectures were explored:

2.3.1 Neural Collaborative Filtering (NCF)

The NCF framework [4] extends matrix factorization by replacing the inner product with a neural architecture to learn nonlinear user-item interactions. After thoroughly studying the NCF paper, we came to understand why NCF is superior to traditional matrix factorization. The comprehensive paper on NCF outlines three key components:

1. Generalized Matrix Factorization (GMF): A neural network generalization of matrix factorization that allows for learning more complex interactions through non-linear transformations. In our initial research stages, we considered implementing GMF as described in the NCF paper. GMF works by transforming categorical data (user and movie IDs) into dense vector representations (embeddings) that capture hidden patterns in user preferences.

The embedding process we studied works as follows:

Consider a dataset with users and movies:

Users	M1	M2	M3
U1	5	3	1
U2	4	2	2
U3	?	2	3

Figure 2: EXAMPLE MATRIX

Where ? represents missing ratings that the model aims to predict.

Matrix factorization in GMF decomposes the original rating matrix into two lower-dimensional matrices: User Embedding Matrix (U) and Movie Embedding Matrix (M). The dot product of these matrices reconstructs the original rating matrix, allowing prediction of missing ratings.

For instance, with user embeddings [0.8, -0.2, 0.5] and movie embeddings [0.9, 0.3, -0.5], the reconstructed rating matrix would be: "‘ Users M1 (0.9) M2 (0.3) M3 (-0.5) U1 (0.8) 0.8 × 0.9 = 0.72 0.8 × 0.3 = 0.24 0.8 × -0.5 = -0.4 U2 (-0.2) -0.2 × 0.9 = -0.18 -0.2 × 0.3 = -0.06 -0.2 × -0.5 = 0.1 U3 (0.5) 0.5 × 0.9 = 0.45 0.5 × 0.3 = 0.15 0.5 × -0.5 = -0.25 "‘

The key innovation in GMF is that it applies a non-linear activation function to this process, allowing for more complex patterns to be captured beyond what traditional matrix factorization can achieve.

2. Multi-Layer Perceptron (MLP): A variant that uses multiple fully connected layers to learn complex mappings between user and item latent factors. Our current implementation focuses on this component of the NCF architecture.

3. Neural Matrix Factorization (NeuMF): A fusion model that combines GMF and MLP to leverage both the linearity of GMF and the non-linearity of MLP, achieving better performance than either component alone.

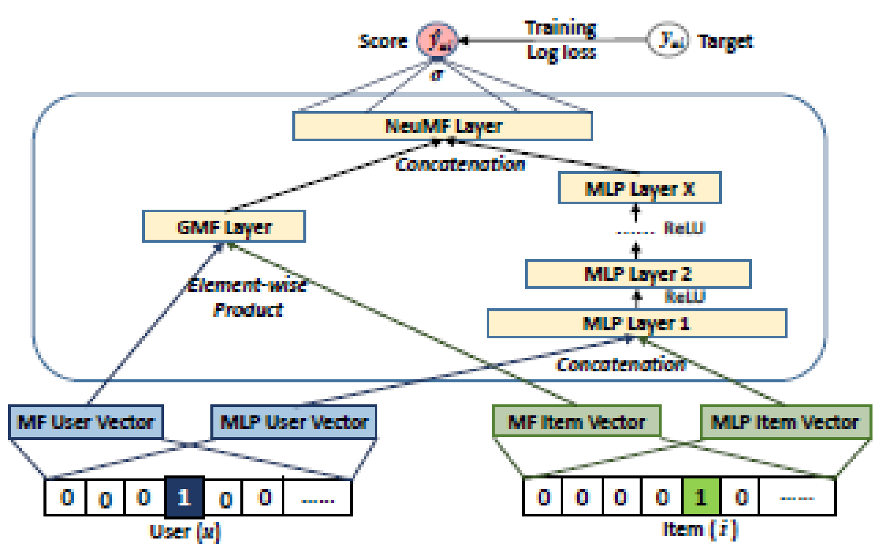


Figure 3: Neural matrix factorization model [4]

NCF demonstrates significant improvements over traditional matrix factorization by capturing complex relationships between user and item latent factors. The study verified that deeper architectures typically yield better performance, highlighting the importance of learning high-order feature interactions.

2.3.2 Sequence-Aware Recommendation Models

Given the temporal nature of user preferences and consumption patterns, we explored sequence-aware recommendation models as a potential direction:

1. Recurrent Neural Networks (RNNs): These networks, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), model sequential dependencies in user behaviors [6]. By capturing the temporal dynamics in user-item interactions, RNNs offer improved performance for session-based recommendations.

However, during our investigation, we ultimately dropped RNNs from consideration as they did not take user ratings into account and relied only on the sequence of watched movies, which limited their recommendation effectiveness. As noted in [2] work on Collaborative Filtering with Recurrent Neural Networks, this approach has limitations when user ratings are a crucial factor in recommendation quality.

2. Transformer Models: Originally introduced for natural language processing tasks [21], transformers have been adapted for sequential recommendation tasks. The self-attention mechanism allows for capturing long-range dependencies without the sequential constraints of RNNs.

3. BERT4Rec: This approach [19] adapts the BERT (Bidirectional Encoder Representations from Transformers) architecture for sequential recommendation. By employing bidirectional self-attention and masked language modeling objectives, BERT4Rec effectively captures contextual information from both historical and future interactions.

We considered applying masking in transformer-based BERT4Rec to predict movies, and then filtering out genres that aligned with current emotions. This would have been a hybrid approach where the sequential model would first generate recommendations based on viewing history, and then a post-processing emotion-based genre filter would refine these recommendations to match the user's current emotional state. However, we ultimately decided against this approach for the same reason we dropped RNNs: these sequential models don't sufficiently account for user ratings and user similarity, which are crucial factors for our recommendation system.

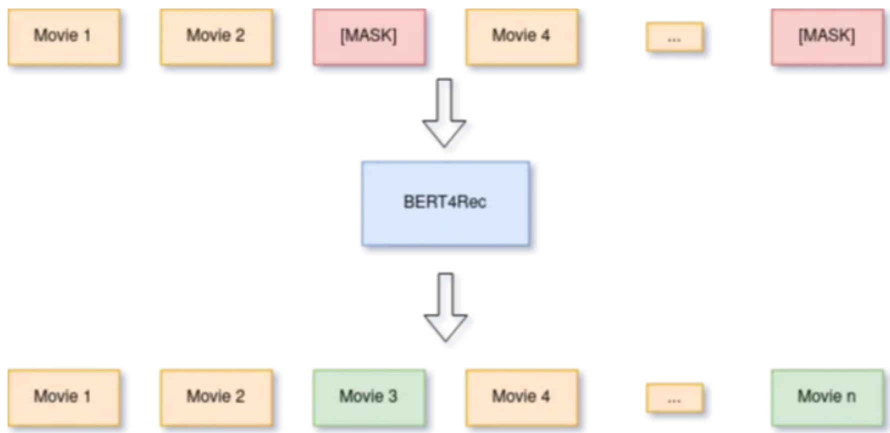


Figure 4: BERT4REC Masking [7]

2.3.3 Autoencoder-Based Approaches

Autoencoders represent another significant neural architecture we explored for recommendation tasks:

1. Vanilla Autoencoders: These neural networks compress input data into a lower-dimensional representation and then reconstruct the original input, learning useful feature representations in the process [16].

2. Denoising Autoencoders: These extend vanilla autoencoders by intentionally corrupting input data and training the network to recover the original uncorrupted input, improving robustness and representation learning [18].

3. Variational Autoencoders (VAEs): These generative models learn a probabilistic latent representation, enabling more expressive modeling of user preferences [9].

During our exploration, we considered an approach that would pass concatenated user and movie matrices through autoencoders to check reconstruction loss during training. This method would involve compressing the interaction matrix into a lower-dimensional latent space via an encoder, capturing essential features in a bottleneck layer, and then reconstructing the original input through a decoder, forming the reconstruction matrix. The autoencoder would function as an unsupervised learning technique in collaborative filtering, with the input matrix also serving as the target matrix.

While this approach showed promise, we ultimately pursued alternative directions better aligned with our project goals, particularly the extension of NCF with emotion-aware components.

2.3.4 Emotion-Aware Recommendation Systems

The integration of emotional context into recommendation systems represents a novel direction that has gained traction in recent years:

1. Emotion Recognition: Several approaches for emotion detection from text have been proposed, including lexicon-based methods and more advanced machine learning techniques [12].

2. Emotion-Genre Correlation: Studies have established links between specific emotions and preferences for particular movie genres [20], suggesting that recommendations aligned with a user's current emotional state might improve satisfaction.

3. Contextual Recommendation: Context-aware recommendation systems incorporate situational factors, including emotional states, to provide more relevant suggestions [1].

Our exploration led us to consider the Go Emotions dataset and BERT-based models for emotion detection as a foundation for emotion-aware recommendations.

2.4 Our Current Implementation

After extensive exploration of the recommendation system landscape, we developed a hybrid approach that combines neural collaborative filtering with emotion-aware recommendations. Key components of our implementation include:

1. Neural Collaborative Filtering with Genre Enhancement: We're currently extending the NCF framework by incorporating genre information into the model architecture. Specifically, we're adding a genre embedding layer to the MLP component of NCF, wherein we concatenate a multi-hot encoded vector representing movie genres to the existing model setup and train the model accordingly. This approach enhances the recommendation quality by leveraging both user-item interaction patterns and content-based features.

2. Emotion Detection Model: A BERT-based multi-label classification model trained on the Go Emotions dataset to identify emotional states from user-provided text.

3. Emotion-Genre Mapping: A mapping framework that associates detected emotions with relevant movie genres, considering both resonance (genres that match the emotional state) and contrast (genres that might counter the emotional state).

4. Recommendation Generation: A system that combines predicted ratings from the NCF model with emotion-derived genre preferences to generate personalized movie recommendations aligned with the user's current emotional state.

2.4.1 Hybrid Nature of Our Current Model

Our current implementation represents a true hybrid approach that combines multiple recommendation paradigms:

1. **Collaborative Filtering Aspect:** Through the NCF model, we incorporate user-item interaction patterns and user similarities, capturing the collaborative filtering principle that similar users have similar preferences.

2. Content-Based Aspect: By integrating genre information directly into the NCF architecture (through multi-hot encoded genre vectors), we incorporate content-based features that allow the model to understand movie characteristics beyond just user interactions.

3. Context-Aware Aspect: The emotion detection component provides contextual awareness by considering the user's current emotional state, allowing for dynamic recommendations that adapt to changing user contexts.

4. Feature Augmentation Strategy: Rather than using a simple ensemble method that combines separate recommendation systems, we employ a feature augmentation strategy by directly integrating genre features into the neural network architecture and using emotion-genre mappings to refine the recommendations.

This hybrid approach addresses several limitations of traditional recommendation systems by: - Incorporating emotional context to provide more personalized recommendations - Leveraging the expressive power of neural networks to model complex user-item relationships - Integrating genre information as an additional signal for recommendation generation - Offering users control over whether recommendations should match their current mood or provide contrast

We are currently working with the MovieLens dataset to implement and evaluate our approach. The implementation involves training a Neural Collaborative Filtering model with our genre enhancements, developing an emotion detection component, and integrating these systems to provide emotion-aware movie recommendations.

Chapter *3*

Proposed Work

From Model Exploration to Final System Our goal was to build an Emotion-Aware Movie Recommendation System that can adapt to users' changing emotional states and suggest suitable content accordingly. This involved designing a dual-model architecture that combines emotion detection from user-provided text with genre-aware neural recommendations trained on user preferences. Our final solution is the result of multiple stages of experimentation, model evaluation, and critical decision-making.

3.1 Content-Based vs. Collaborative Filtering

Selecting the Correct Foundation: We started our project with a comparison between Content-Based Filtering and Collaborative Filtering, the two traditional methods in recommender systems.

Decision: We opted for Collaborative Filtering because of its better scalability and capacity to model user- item relationships without demanding extensive movie metadata.

3.2 Alternatives Explored and Why They Were Dropped

As part of improving CF, we explored several deep learning-based extensions

3.2.1 Autoencoders

We tried using stacked autoencoders to learn patterns in how users interact with movies. But autoencoders mainly focus on rebuilding the original data without using labels, so they didn't work well for our goal of including extra information like emotions and genres.

3.2.2 Transformers (e.g., BERT4Rec)

Transformers are the latest for sequential recommendations. But as we weren't modeling temporal consumption behaviors, and didn't have ordered session data, we decided against the additional complexity they came with .

3.2.3 Matrix Factorization

Humanize AIWe also used classical MF-based collaborative filtering, which decomposes the rating matrix into latent item and user features. Yet, MF is shallow and linear, incapable of modeling high-order, complex user-movie interactions or processing categorical data such as genre in an efficient manner .

3.3 Final Architecture

Dual-Model Design After critical evaluation, we finalized a two-part system:

A. Emotion Detection Model (Model 1): Text input from the user (e.g., ”I’m feeling nostalgic today”).

Approach: We fine-tuned a BERT-based classifier on the GoEmotions dataset, which has more than 58,000 Reddit comments annotated with 27 subtle emotions and neutral.

Why GoEmotions? It offers a large emotional lexicon much richer than simple sentiment, allowing the system to pick up on subtle user moods such as ’nostalgia’, ’excitement’, or ’disappointment’. A softmax score vector for emotion labels. The highest predicted emotion is then projected to a set of resonant and contrasting genres through a carefully curated Emotion-Genre Mapping Table.

User Control: We designed an interactive layer where, upon emotion detection, the user is prompted: ”Would you like to watch something that fits your current mood or something to transform it?” Based on which either resonant or contrast genres are picked up for recommendation pipeline.

B. Movie Recommendation Model (Model 2 - NCF) Input Features:

- **User ID** (learned embedding)
- **Movie ID** (learned embedding)
- **Genre** (multi-hot encoded vector from mapped emotion)

Model Used: Neural Collaborative Filtering (NCF), deep learning-based collaborative filtering technique that generalizes the old MF by learning non-linear user-item interactions.

Architecture:

We employ an MLP (Multi-Layer Perceptron) instead of the matrix dot product. The MLP processes the concatenation of user and item embeddings, and genre information through several hidden layers. Each hidden layer learns high-order latent factors between users, items, and genres.

Output layer outputs a user-specific rating for every movie.

Why MLP in NCF? MLP learns non-linear interaction patterns, in contrast to the conventional matrix factor- ization which only simulates simple dot products. This equips the system with the capability to recognize more sophisticated relationships, like a user’s mood affecting preferences differently in different genres.

C. Merging the Two Models The final pipeline integrates both models:

Combining the Two Models The last pipeline combines both models:

User inputs text → Emotion Detection (BERT-GoEmotions) → Emotion Label

Emotion Label → Emotion-to-Genre Mapping → Resonant/Contrast Genres

Genre Info + UserID + MovieID → NCF Model → Top-K Movie Recommendations

This segregation makes it possible to train every module separately and update modularly in the future.

3.4 Evaluation Strategy We intend to measure the performance of our final model based on the following metrics:

• **Hit Ratio@K (HR@K):** Checks if the ground truth movie is among the top-K recommendations.

• **Normalized Discounted Cumulative Gain (NDCG@K):** Checks the ranking quality — greater scores for correct predictions that occur higher up in the list.

These metrics will assist us in deciding not just whether the appropriate film was suggested, but also how well it ranked versus less relevant ones .

3.4 Technical Background

Over the past few years, Deep Learning (DL) methods have seen significant improvement in numerous applications such as natural language processing (NLP), recommendation systems, and emotion detection. Hand-crafted features or shallow models are generally used by conventional machine learning techniques, but deep learning-based models like transformers and neural collaborative filtering (NCF) provide proven handling of high-dimensional intricate data.

For our project, DL methods were essential because of the difficulty involved in emotion detection from text and producing customized movie suggestions. Specifically, two significant DL methods were applied:

- **Transformers:**Transformers are used for emotion detection and are optimized for dealing with sequential data such as text. The BERT (Bidirectional Encoder Representations from Transformers) model was selected because its performance in numerous NLP tasks, including emotion detection, is state-of-the-art.
- **Neural Collaborative Filtering (NCF):**Used in the recommendation system, NCF combines collaborative filtering and neural networks, learning the user and movie embeddings and utilizing the embeddings to predict ratings or make recommendations.

1. Emotion Detection Code (BERT-based Model)

The emotion detection model is the initial crucial element of our system, which uses a BERT-based model fine-tuned on the GoEmotions dataset to identify emotions in user input text.

Key Techniques:

- BERT is a pre-trained trans-former model that is used to process and comprehend contextual relations in text. BERT captures both the left and right context of a word in a sentence (hence bidirectional) by employing attention mechanisms, making it strong to be used on tasks such as emotion detection.

- Multi-label Classification: Since every text could have more than a single emotion, the model is structured to predict various emotions (multi-label classification) and not an individual emotion.

WHY BERT?

Approaches we have followed that lead to BERT: For an emotion-aware movie recommendation system, emotion detection is crucial to ensure the system knows about the emotional state of the user and correctly maps it to corresponding movie genres. BERT (Bidirectional Encoder Representations from Transformers) is the most solid model for this purpose and performs better than other approaches in many important aspects.

1) Traditional NLP Techniquess Sentiment Analysis (TextBlob, VADER, AFINN)

Although these classic approaches label text as positive, negative, or neutral, they do not catch the nuance of emotions like joy, sadness, anger, or surprise. Their dependence on pre-specified lexicons and rule-based methods makes them miss subtle emotional expressions like sarcasm or mixed feelings.

Why BERT is Better: BERT's contextualized, deep understanding of language enables it to capture the subtle feelings in text by comparing words with the context they occur in. Contrary to rule-based approaches, BERT has the capability of recognizing subtle emotional signals that could occur under varied contexts, for instance, capturing "fear" within a joke context which would otherwise go unnoticed through other methods .

2) Machine Learning-Based Approaches

Classical ML Models (SVM, Random Forest, Naive Bayes)

These models are generally based on heavy feature engineering (e.g., TF-IDF or word embeddings) to map text to a numerical representation. Additionally, they depend on hand-tuned, labeled datasets in order to enhance performance.

Why BERT is Better: BERT does away with the necessity for hand-engineered features. Pretrained on huge text corpora, it discovers rich, high-level language representations, which are readily fine-tuned for emotion detection tasks (e.g., utilizing GoEmotions). This greatly minimizes the labor for feature extraction without compromising on higher performance in recognizing intricate emotional states.

Deep Learning Models (LSTMs, CNNs)

Although LSTMs and CNNs are able to learn some emotional patterns, they rely on static word embeddings (such as Word2Vec or GloVe), which do not catch context as well as BERT's bidirectional model. Also, these models tend to need lots of data and are vulnerable to overfitting when data is limited.

Why BERT is Better: BERT's transformer-based model processes text bidirectionally, catching the entire context of words both from the left and right. This allows BERT to better identify more complex relationships between text than LSTMs or CNNs, which normally read text in a single direction.

3) Neural Network-Based Approaches

Transformers (BERT, DistilBERT, RoBERTa)

Among transformer-based models, BERT is now the model of choice for a wide range of NLP applications, including emotion detection. Its bidirectional contextualization of text allows it to pick up subtle emotional cues better than other models, such as LSTM and CNN, can match.

Why BERT Stands Out: BERT's capability of processing text from both directions creates a more rich un-derstanding of context, essential to detect emotions accurately. Fine-tuning BERT on emotion specific data like GoEmotions (which tags 27 unique emotions) enables it to focus on this task,beating most other models at emotion detection.

4) Large Language Models (LLMs) APIs

OpenAI's GPT Models (e.g., GPT-4, GPT-3.5)

Although GPT models are capable of emotion detection as well, they are geared more towards text generation than classification tasks. Furthermore, employing GPT models via APIs adds latency and recurring costs, particularly in real-time usage.

Why BERT is Better: BERT's fine-tuning on emotion datasets such as GoEmotions makes it even more specialized in emotion detection. It provides better accuracy than GPT models for this purpose.

In addition, BERT can be used offline, achieving affordable, scalable solutions for emo-

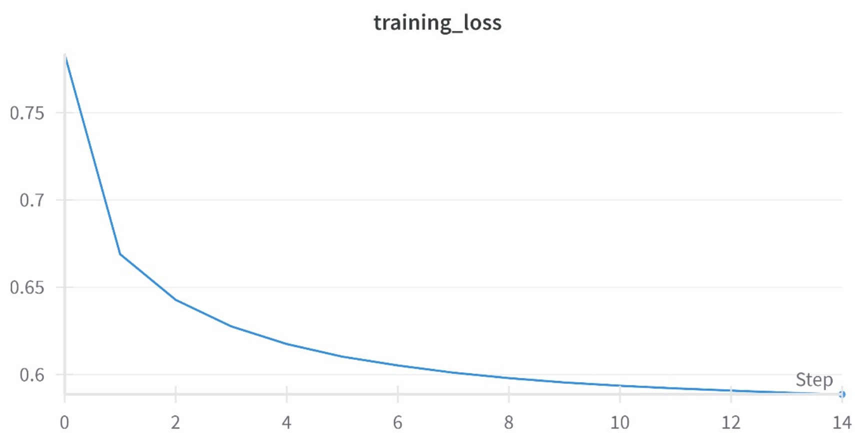
tion detection in real-time without the GPT API overhead or latency. .

In summary, BERT's sophisticated contextual intelligence, pretrained performance, and high performance in emotion detection qualify it as the perfect model for your emotion-sensitive movie recommendation system. With the application of BERT in emotion detection, you are able to make sure that the system precisely identifies and understands user emotions, hence providing more relevant and personalized movie recommendations.

Chapter *4*

Simulation and Results

4.1 Results



Recommendation model training loss

Recommendation model evaluation loss

Epoch	Training Loss	Validation Loss
1	0.091300	0.086535
2	0.075700	0.082690
3	0.054100	0.087214

Emotion model loss

4.2 Simulation

```
ENTER A SENTENCE-I was having a nice day until I slipped on a banana peel and now I am pissed.
Would you like to match your mood, shift it, or keep it neutral? (match/shift/neutral): shift

==== Mood-Based Movie Recommender ====

Loading movie data from: ./filtered_data.csv
Loaded 32000204 ratings, 200948 users, 84432 movies
Loading emotion detection model...
Loading movie recommendation model...
Analyzing mood from text: 'I was having a nice day until I slipped on a banana peel and now I am pissed.'

Detected emotions:
  - annoyance: 0.668
```

```
Relevant movie genres for your mood:
  - Music: 0.304
  - Animation: 0.304
  - Comedy: 0.130
  - Crime: 0.130
  - Drama: 0.130
DEBUG: Selected top genres for one-hot encoding: [('Music',
np.float32(0.13043478)), ('Crime', np.float32(0.13043478))]

Generating recommendations for user 4 based on mood...

✨ Top Movie Recommendations Based on Your Mood ✨

1. John Mulaney: New In Town (2012)
   Genres: Comedy
   Predicted Rating: 4.41/5.0

2. There Once Was a Dog (1982)
   Genres: Animation|Children|Comedy
   Predicted Rating: 4.40/5.0

3. James Acaster: Cold Lasagne Hate Myself 1999 (2020)
   Genres: Comedy
   Predicted Rating: 4.39/5.0
```

```
emotion_to_genres = {
    "joy": {
        "resonant": ["Comedy", "Family", "Animation", "Music"],
        "contrast": ["Horror", "War", "Thriller"]
    },
    "amusement": {
        "resonant": ["Comedy", "Animation", "Family"],
        "contrast": ["Film-Noir", "Crime", "War"]
    },
    "excitement": {
        "resonant": ["Action", "Adventure", "Sci-Fi", "Fantasy"],
        "contrast": ["Documentary", "Drama"]
    },
    "optimism": {
        "resonant": ["Adventure", "Romance", "Family"],
        "contrast": ["Thriller", "War", "Horror"]
    },
    "pride": {
        "resonant": ["Biography", "History", "Drama"],
        "contrast": ["Horror", "Film-Noir"]
    },
    "gratitude": {
        "resonant": ["Romance", "Family", "Comedy"],
        "contrast": ["Horror", "War"]
    },
    "love": {
        "resonant": ["Romance", "Drama", "Music"],
        "contrast": ["Action", "Horror"]
```

Emotion to Genre Mapping Example

Chapter *5*

5.1 Conclusions

This project shows the potential and usefulness of incorporating emotion awareness into film recommendation systems. Using a transformer-based model for detecting emotions from user text and translating these emotions to film genres, we facilitate contextually appropriate suggestions that either match or shift the user's mood. Integrated with Neural Collaborative Filtering, this strategy strikes a balance between past user behavior and live emotional context, providing a more personalized and emotionally aware recommendation experience.

5.2 Scope of further work

5.2.1 What We Will Do

To expand and polish the current system, we plan to implement the following:

- **User Interface Creation**: We will create an interactive front-end through which users can enter free-text emotions and be provided with movie recommendations in a user-friendly interface.
- **Feedback Mechanism**:Feedback will be included through which users can rate or respond to the recommended movies, based on feedback from users to refine future suggestions.
- **Group Emotion Detection**: Our intention is to implement group-based emotion aggregation functionality, where emotional states of more than one user are processed to recommend movies appropriate for the group mood.

5.2.2 What Can Be Done

A number of advanced features can be considered in future versions of the system to enhance its capabilities and emotional intelligence:

- **Personalized Emotion-to-Genre Mapping**: Substitute fixed mappings with dynamic, user-dependent links between emotions and genres from historical interactions and feedback.

- **Temporal Emotion Tracking**: Track and analyze emotional patterns over time to offer recommendations that correspond to users' longer-term mood trends or emotional objectives.
- **Hybrid Recommendation Approaches**: Combine content-based filtering and knowledge graphs with collaborative filtering to diversify and personalize recommendations better.
- **Emotion-Based Filtering and Search**:Enable users to search or filter films based on desired emotional responses (e.g., ”uplifting comedies” or ”relaxing and nostalgic dramas”).
- **Emotion-Augmented Explanations**: Provide transparent, explainable rationales for recommendations, merging collaborative understanding and emotional context (e.g., ”Recommended for your current relaxed state and film preferences”).

Bibliography

[1] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems (TOIS)*, 23(1):103–145, 2005.

[2] R. Devooght and H. Bersini. Collaborative filtering with recurrent neural networks. *Neural Networks*, 2017.

[3] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins University Press, 3rd edition, 1996.

[4] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua. Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web*, pages 173–182, 2017.

[5] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53, 2004.

[6] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk. Session-based recommendations with recurrent neural networks. In *International Conference on Learning Representations (ICLR)*, 2016.

[7] S. Jain. Build your own movie recommender system using bert4rec. `https://medium.com/data-science/build-your-own-movie-recommender-system-using-bert4rec-92e4e34938c5`, 2023. Accessed: 2025-04-25.

[8] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[9] D. Liang, R. G. Krishnan, M. D. Hoffman, and T. Jebara. Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 World Wide Web Conference*, pages 689–698, 2018.

[10] P. Lops, M. De Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*, pages 73–105. Springer, Boston, MA, 2011.

[11] A. Mnih and R. R. Salakhutdinov. Probabilistic matrix factorization. In *Advances in neural information processing systems*, pages 1257–1264, 2008.

[12] S. M. Mohammad and P. D. Turney. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465, 2013.

[13] S. Robertson. Understanding inverse document frequency: on theoretical arguments for idf. *Journal of Documentation*, 60(5):503–520, 2004.

[14] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pages 285–295, 2001.

[15] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen. Collaborative filtering recommender systems. In *The adaptive web*, pages 291–324. Springer, Berlin, Heidelberg, 2007.

[16] S. Sedhain, A. K. Menon, S. Sanner, and L. Xie. Autorec: Autoencoders meet collaborative filtering. In *Proceedings of the 24th International Conference on World Wide Web*, pages 111–112, 2015.

[17] A. Singhal. Modern information retrieval: A brief overview. *IEEE Data Engineering Bulletin*, 24(4):35–43, 2001.

[18] F. Strub and J. Mary. Collaborative filtering with stacked denoising autoencoders and sparse inputs. In *NIPS workshop on machine learning for eCommerce*, 2015.

[19] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, and P. Jiang. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 1441–1450, 2019.

[20] M. Tkalčič, A. Kosir, and J. Tasic. Affective recommender systems: The role of emotions in recommender systems. In *Proceedings of the RecSys 2011 Workshop on Human Decision Making in Recommender Systems*, pages 9–13, 2011.

[21] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

[22] Y. Zhou, D. Wilkinson, R. Schreiber, and R. Pan. Large-scale parallel collaborative filtering for the netflix prize. In *International conference on algorithmic applications in management*, pages 337–348. Springer, Berlin, Heidelberg, 2008.