International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Music Recommendation System on Spotify Using Deep Learning

{BE Students}* Pranali Balkrushna Chopade, Nikita Baban Chormale,

Pratiksha Pratap Shinde, Vishakha Suryakant Shitole.

{Faculty} Sachin Sukhadev Bere, Dinesh Bhagwan Hanchate**

Department of Computer Engineering, Dattakala Group of Institutions Faculty of Engineering, Swami-Chincholi Bhigwan-413130, University of Pune, Maharashtra, INDIA.

vishakhashitole19@gmail.com

Abstract

In the era of personalized digital experiences, music recommendation systems play a vital role in enhancing user satisfaction by delivering contextually relevant content. This project presents a deep learning-based music recommendation system integrated with Spotify that tailors music suggestions based on user age and current weather conditions. The system leverages computer vision techniques for age detection through facial analysis using Convolutional Neural Networks (CNNs), while weather data is retrieved from an external API based on the user's location. By combining these inputs, the system dynamically curates playlists that match both the user's demographic profile and the environmental context. The model is trained on a diverse dataset of facial images for accurate age classification and integrates seamlessly with Spotify's Web API for real-time music playback. Experimental results demonstrate improved user engagement and satisfaction through personalized music experiences. This approach highlights the potential of deep learning in enhancing recommender systems by incorporating both human and environmental factors.

1. Introduction

In today's digital age, music streaming platforms like Spotify have revolutionized how people discover and consume music. With millions of tracks available, helping users find content that aligns with their tastes has become a critical feature. Music recommendation systems address this challenge by leveraging user behavior, song attributes, and listening patterns to suggest relevant tracks.

Spotify's recommendation engine is a sophisticated system that combines collaborative filtering, content-based filtering, and deep learning techniques. Deep learning, in particular, has enabled more accurate and personalized recommendations by modeling complex patterns in large-scale user interaction data. Neural networks can learn high-level representations of songs and users, capturing subtle similarities in sound, mood, and genre that traditional methods often miss.

This project explores the development of a music recommendation system using deep learning models. It involves analyzing user listening histories, audio features (such as tempo, energy, and danceability), and other metadata to train a neural network capable of predicting and recommending songs. The goal is to improve personalization and user satisfaction on platforms like Spotify by providing smarter and more dynamic music suggestions.

2. The Role of Deep Learning in Spotify's Music Recommendation System

Deep learning has significantly enhanced the performance and personalization of music recommendation systems by enabling the modeling of complex patterns in user behavior and audio content. Traditional recommendation techniques, such as collaborative filtering and content-based filtering, have limitations in handling sparse data, cold-start problems, and capturing non-linear relationships. Deep learning overcomes many of these challenges through its ability to learn rich, high-dimensional representations.

Here are the key roles deep learning plays in music recommendation:

1. Audio Feature Extraction

Deep learning models, particularly Convolutional Neural Networks (CNNs), are effective at analyzing raw audio signals or spectrograms to extract meaningful features. Unlike manual feature engineering, CNNs automatically learn patterns such as rhythm, pitch, timbre, and genre, which are crucial for content-based recommendation.

2. User Behavior Modeling

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are used to model sequential user behavior over time. They help capture listening habits, transitions between songs, and temporal preferences, allowing for more context-aware recommendations.

3. Latent Representation Learning

Autoencoders and embedding layers are employed to learn dense vector representations (embeddings) of users and songs. These embeddings capture latent similarities that are not immediately obvious, improving the quality of recommendations even with sparse interaction data.

4. Hybrid Recommendation Models

Deep learning enables the integration of multiple data sources—such as user interaction history, audio features, lyrics, and metadata—into a unified model. This hybrid approach improves robustness and recommendation accuracy, especially for new or less popular songs.

5. Personalization and Context Awareness

Deep learning models can consider contextual information like time of day, mood, device type, or location. Using models like Deep Neural Networks (DNNs) and attention mechanisms, the system adapts recommendations in real time to suit the user's current context.

6. Scalability

Modern deep learning architectures, when trained on large datasets, can scale effectively to millions of users and songs. This makes them ideal for platforms like Spotify, which require fast and accurate recommendations across vast catalogs.

4. Deep Learning in Spotify's Recommendation System

Deep Learning in Spotify's Recommendation System

Spotify leverages deep learning as a core component of its recommendation engine to deliver personalized music experiences to millions of users worldwide. With over 100 million tracks and diverse user preferences, traditional recommendation approaches alone are insufficient. Deep learning enhances Spotify's ability to model both content and user behavior, enabling smarter, more context-aware suggestions.

1. Audio Analysis with CNNs

Spotify uses Convolutional Neural Networks (CNNs) to analyze raw audio signals and generate high-level representations of songs. These models learn patterns such as genre, tempo, instrumentation, and mood directly from spectrograms, allowing for content-based recommendations even when user interaction data is limited (i.e., cold-start songs).

2. User Behavior Modeling with RNNs

To model the sequence of user interactions, Spotify employs Recurrent Neural Networks (RNNs) and their variants (e.g., LSTMs, GRUs). These networks capture listening patterns over time, such as daily habits or transitions between music genres, which improves session-based recommendations and playlist generation.

3. Collaborative Filtering with Embeddings

Deep learning models use embedding layers to represent users and tracks in a shared latent space. This approach allows the system to compute similarity scores more effectively, even for users or tracks with sparse data, enhancing collaborative filtering capabilities.

4. Hybrid Deep Learning Models

Spotify combines audio content, user behavior, and metadata (e.g., genre, artist, release year) in hybrid deep learning models. This integration helps produce more holistic recommendations that align with both user tastes and content characteristics.

5. Playlists and Discover Weekly

One of Spotify's most popular features, *Discover Weekly*, uses deep learning to generate personalized playlists based on a user's listening history and similar users' preferences. The system blends collaborative and content-based filtering with neural models to predict songs the user might enjoy but hasn't heard yet.

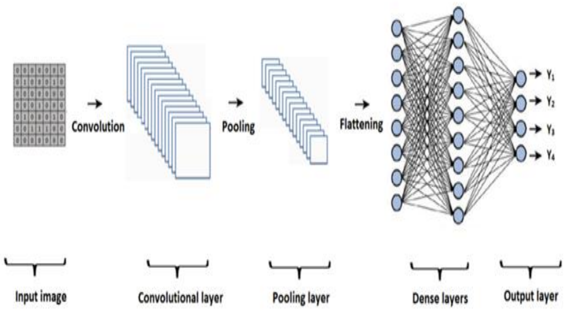
6. Context-Aware Recommendations

Deep learning also powers Spotify's contextual recommendation features. By incorporating factors such as time of day, device type, or current activity (e.g., workout, relaxation), the system can adapt music recommendations dynamically, enhancing the user experience.

- ## Algorithms:

A Convolutional Neural Network (CNN) is a specialized deep learning architecture designed to process and analyze visual data, such as images and videos.

Inspired by the human visual system, CNNs excel at detecting patterns and structures in visual inputs, making them foundational in tasks like image classification, object detection, and facial recognition.



- **Steps in CNN :**

1. **Input Layer** : Receives the raw pixel values of the image (e.g., 28×28×1 for grayscale). The input is usually an image represented as a 3D matrix (height × width × channels). o Example: A color image of size 64×64 has dimensions 64×64×3 (R, G, B channels)
2. **Convolutional Layer :** Applies a set of filters (kernels) to extract features like edges, textures, etc. Produces feature maps. This is the core building block of a CNN. It uses filters/kernels (small matrices like 3×3 or 5×5) that slide over the input image and perform element-wise multiplication and summation. This operation extracts features like edges, textures, or shapes.
3. **ReLU Layer (Activation) :** Applies non-linearity to the feature maps. Converts negative values to zero.
4. **Pooling Layer :** Down samples the feature maps. Common types: Max Pooling, Aver age Pooling.

Context-Aware Recommendation Engine

Spotify's engine incorporates real-time and inferred contextual data, such as time, location, activity, device, age group, and even weather conditions, to recommend content that fits the moment. Deep learning helps model these contexts effectively.

How CNN Works in Age Detection:

1. Input: Facial image.
2. Preprocessing: Face is detected and cropped. Normalization and resizing (e.g., 224×224 RGB image).
3. Convolutional Layers: Extract facial features like wrinkles, skin texture, eye bags, etc.
4. Pooling Layers: Reduce spatial size while retaining key features.
5. Fully Connected Layers: Use extracted features to predict age.
6. Output: Regression (exact age) or classification (age group like 0–10, 11–20, etc.)

How CNN works in Weather Detection :

1. Live Metadata: GPS location (longitude, latitude)Temperature (via API like OpenWeatherMap)
2. CNN Feature Extraction: Cloud coverage Brightness/sunlight Raindrops, snowflakes, fog
3. Metadata Fusion: Non-image data (location, temperature, season) is fed into a fully connected (dense) neural network layer.
4. Final Prediction Output layer classifies the current weather as: Sunny 🌞, Cloudy ☁️, Rainy 🌧, Snowy ❄️, Foggy.

5. Key Components of Spotify's Recommendation System

1. Age Detection Module

While Spotify doesn't directly ask users for their age in all markets, age can be **inferred** using:

- **User Profile Data**: If age or birth year is provided during signup.
- **Behavioral Patterns**: Listening history can hint at a user's age group (e.g., 90s hits vs. modern pop).
- **Deep Learning Classification Models**: Neural networks trained on user behavior and demographics to predict age brackets.
- **Collaborative Filtering Signals**: Age-inferred groups can be clustered to recommend songs popular among similar users.

2. Weather Detection Module

Spotify can integrate real-time weather data through:

- **Location-Based Weather APIs**: If location access is enabled, Spotify can fetch the current weather using external services (e.g., OpenWeatherMap, WeatherStack).
- **Edge Computing & Sensors (on mobile)**: Some mobile devices provide environmental data like temperature or humidity indirectly.
- **Context Tagging for Recommendations**:
  - **Rainy** → Chill, acoustic, lo-fi
  - **Sunny** → Upbeat, energetic, dance
  - **Cold/Snowy** → Warm jazz, mellow songs
  - **Hot** → Tropical, beach-vibe music

Deep learning models can incorporate weather as an **input feature** to recommend songs suited to that emotional or atmospheric state.

3. Multimodal Deep Learning Models

Spotify uses models that take **multiple types of input simultaneously**:

- Audio features (e.g., tempo, mood, danceability)
- User behavior sequences (e.g., recent play history)

4. Dynamic Playlist Generation

Spotify can dynamically curate playlists (e.g., *Mood Booster*, *Rainy Day*, *Chill Hits*) using:

- User's age segment (young adults vs. older adults)
- Local weather conditions

6. Flowchart:

Input Data

↓

Convolution Layer

↓

ReLU Activation Function

↓

Pooling Layer

↓

Flatten Layer

↓

Fully Connected Layer

↓

Output Layer

↓

Model Trained

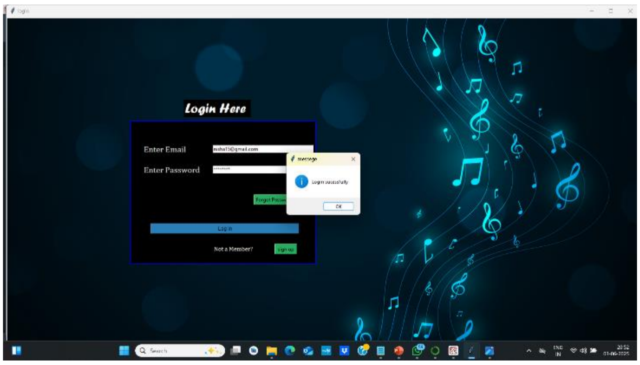
7.Future Trends in Music Recommendation Systems

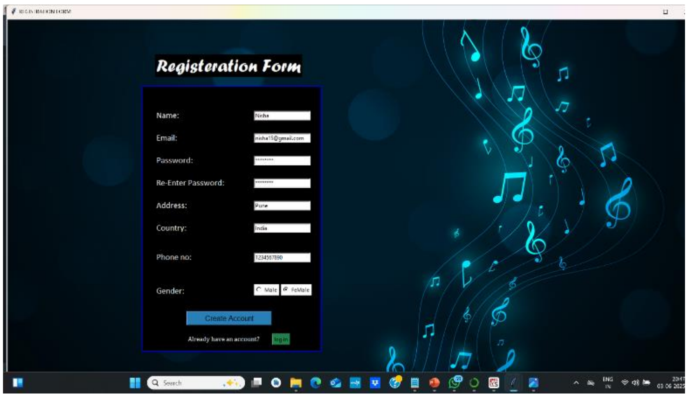
Highlight upcoming research trends and advancements in deep learning that could further improve music recommendation systems.

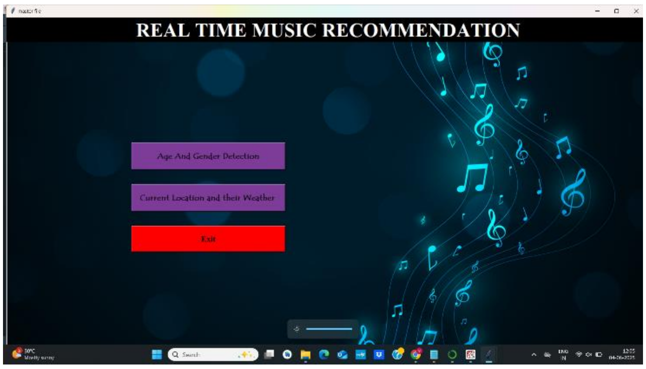
- **Graph Neural Networks (GNNs)**: Emerging trend in modeling user-item relationships through graph structures.
- **Reinforcement Learning**: Optimizing long-term user engagement using reinforcement learning.
- **Transfer Learning**: Applying models trained on one platform to another with minimal retraining.
- **Fairness and Diversity in Recommendations**: How to ensure recommendations are not biased and encourage music discovery.
- **Explainable AI (XAI)**: Making recommendation systems more transparent and interpretable.
- **Reinforcement Learning**: Possible future use of reinforcement learning to better adapt to user preferences in real-time.
- **Personalization Beyond Music**: Integrating recommendations for podcasts, audiobooks, and more into the same framework.
- **Multimodal Learning**: Combining different types of data (audio, lyrics, social) for improved recommendations.
- **Cross-Domain Recommendations**: Recommending not just music, but events, books, or movies based on music preferences.

Result

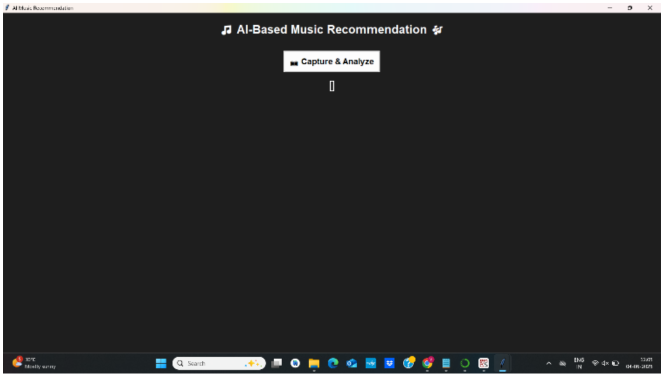

REAL TIME MUSIC RECOMMENDATION
Login
Register
Exit

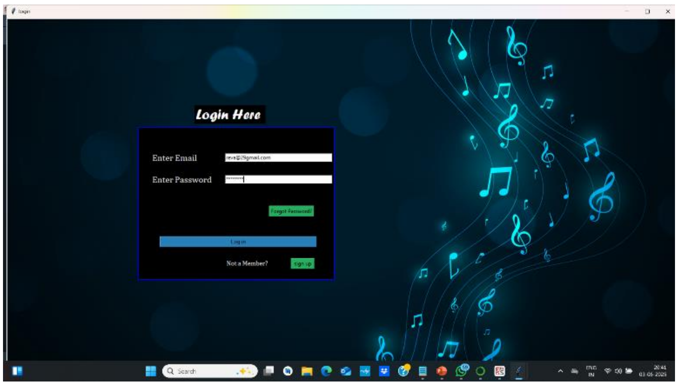

Login Here
Enter Email
Enter Password
Not a Member?

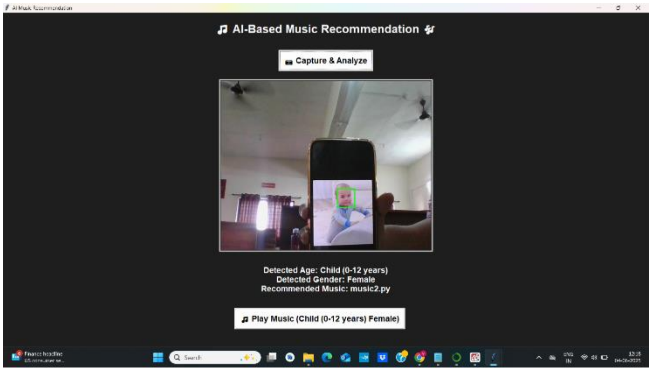

Registration Form
Name:
Email:
Password:
Re-Enter Password:
Address:
Country:
Phone no:
Gender:
Create Account
Already have an account?

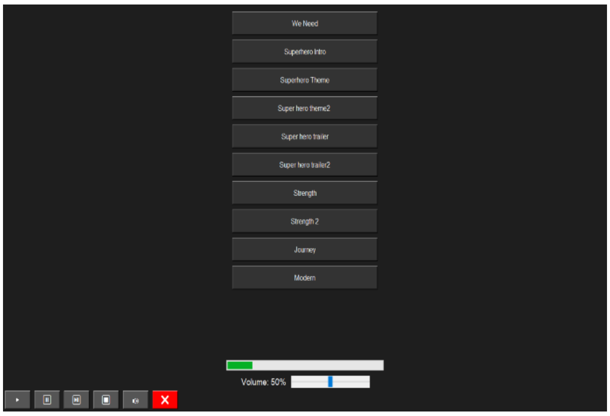

REAL TIME MUSIC RECOMMENDATION
Age And Gender Detection
Current Location and their Weather
Exit

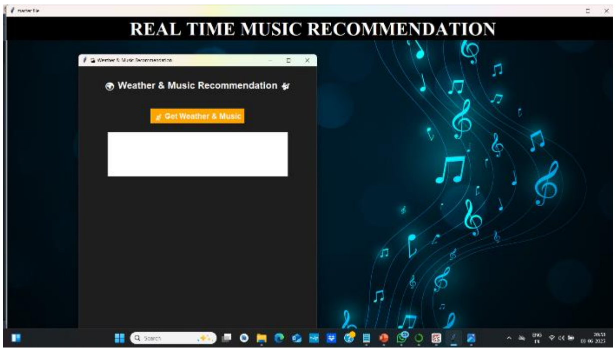

Registration Form
Name:
Email:
Password:
Re-Enter Password:
Address:
Country:
Phone no:
Gender:
Create Account
Already have an account?

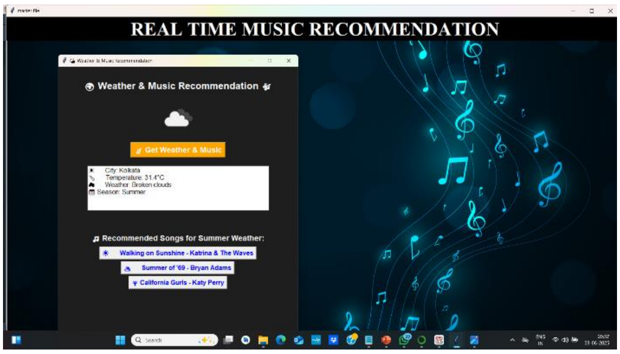

AI-Based Music Recommendation
Capture & Analyze

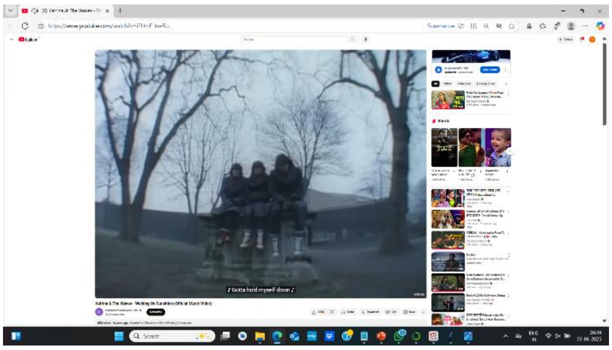

Login Here
Enter Email
Enter Password
Not a Member?


AI-Based Music Recommendation
Capture & Analyze
Detected Age: Child (0-12 years)
Detected Gender: Female
Recommended Music: music2.py
Play Music (Child (0-12 years) Female)









Conclusion

Summarize how deep learning has transformed Spotify's music recommendation system, enhancing user experience by providing more personalized music suggestions. Discuss the potential for further research and innovations.

- **Summary of Deep Learning's Impact**: Highlighting the success of deep learning in Spotify's recommendation system.
- **Future Potential**: The role of ongoing research and new technologies in further enhancing music recommendation systems.

Acknowledgements

This paper would not have been written without the support and encouragement of Dr. D.B.Hanchate, guide of BE Dissertation work. Author's special thanks go to all the professors of computer engineering department of DGOI FOE

Swami chinchholi, for their guidance and for giving her an opportunity to work on Music Recommendation System on Spotify using Deep Learning .

References

- Allawadi, K., & Vij, C. (2023). A Smart Spotify Assistance and Recommendation System. 2023 International Conference on

Advancement in Computation & Computer

Technologies (InCACCT), 286–291.

https://doi.org/10.1109/InCACCT57535.2023.10141810

- Gulmatico, J. S., Susa, J. A. B., Malbog, M. A. F., Acoba, A., Nipas, M. D., & Mindoro, J. N. (2022). SpotiPred: A machine

learning approach prediction of Spotify music popularity by audio features. 2022 Second International Conference on

Power, Control and Computing Technologies (ICPC2T), 1–5.

- Kaminski, G., Dridi, S., Graff, C., & Gentaz, E. (2009). Human ability to detect kinship in strangers' faces: effects of the degree

of relatedness. Proceedings of the Royal Society B: Biological Sciences, 276(1670), 3193–3200.

- Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In The adaptive web: methods and strategies of web

personalization (pp. 325–341). Springer.

- Schildt, H. A., Maula, M. V. J., & Keil, T. (2005). Explorative and exploitative learning from external corporate ventures.

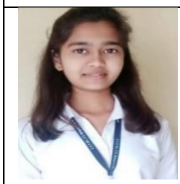
Entrepreneurship Theory and Practice, 29(4), 493–515.

Sedgwick, P. (2012). Pearson's correlation coefficient. Bmj, 345.

- Shannon, C. E. (1948). A mathematical theory of communication. The Bell System Technical Journal, 27(3), 379–423.
- He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 173-182).van den Oord, A., Dieleman, S., & Schrauwen, B. (2013). Deep content-based music recommendation. In *Advances in Neural Information Processing Systems* (pp. 2643-2651).

- **Author Biblography:**

	**Pranali Balkrushna Chopade** Received diploma in computer technology from shriram institute of polytechnic and engineering college paniv. Pursuing BE in computers engineering from Dattakalla group of institutions faculty of engineering certification in web development
	**Nikita Baban Chormale** Received Diploma in computer Technology From Shivnagar Vidya Prasarak Mandal's Institute Of Technology And Engineering Malegaon (BK)-Baramati Malegaon. Pursuing BE in Computer Engineering from

	Dattakala Group of Institutions Faculty of Engineering,Certification in Robotics Process Automation, Full Stack Devlopment with MERN stack.
	**Pratiksha Pratap Shinde** Received Diploma in Computer Technology from Government Polytechnic , Beed. Pursuing BE in Computer Engineering from Dattakala Group of Institutions Faculty of Engineering,Certification in Robotics Process Automation, Full Stack Devlopment with MERN stack, Digital marketing.
	**Vishakha Suryakant Shitole** Received Diploma in Information Technology from SVPM's ITE Malegaon (BK). Pursuing BE in Computer Engineering from DGOIFOE Swami Chincholi Bhigwan. Programming Fundamental's using Python Infosys Springboard Internship 5th Edition Cybersecurity Tech Mahindra Foundation through Skill India Digital Hub Web Development (HTML, CSS, Bootsrap, DBMS, OOP, PHP, React js) Internshala Scholiverse Educare Pvt, Ltd. Introduction to Artificial Intelligence Great Leraning Web Development with Angular & Webpack Mind Luster

	**Sachin Sukhadev Bere** Dr. S.S Bere working as Associate Professor in Dattakala Group of Institutions Faculty of Engineering Bhigwan. He completed his Ph.D. in Computer Science and Engineering, from the, Shri Jagdishprasad Jhabarmal Tibrewala University, Rajasthan India. And completed his M.Tech (CSE) with First class & Distinction from JNTU- Hyderabad affiliated college. He has 14 years of teaching experience, 7 years of Research experience. Presently he isworking as an Associate Professor in Dattakala Group of Institutions Faculty of Engineering Bhigwan, Maharashtra. He published almost 15 research articles in reputed journals and conferences. His interesting areas are Machine Learning, Artificial Intelligence, Deep Learning Techniques and Programming Languages.
	**Dinesh Bhagwan Hanchate** Birth Place: Solapur, received B.E. Computer Walchand College from of Engineering, Sangli (1995), Lecturer in Gangamai College Of Engineering, Dhule (1995-96), Lecturer in S.S.V.P.S. B.S.D. College Of Engineering, Dhule In

	Computer & IT deptt (1996 2005), M.Tech. Computer fromDr.Babasaheb Ambedkar Technological University, Lonere(2002-05), Currently Asst. Prof. Computer Engineering, former H.O.D. (Computer & IT) in Vidya pratishthan's College Of Engineering, Baramati, currently doing research at SGGS Institute Technology Nanded, and affiliated of Engg, to SRTMU, Nanded under the guidance of Dr. Bichkar R.S. G.H. Raisoni College Of Engineering & Management, Wagholi, Pune. Email: dineshbhanchate@gmail.com