RISE RESEARCH INNOVATION SCHOLARSHIP ENTREPRENEURSHIP

DS 4420: Machine Learning and Data Mining 2
Dr. Gerber
2 Dec. 2025

Northeastern University

How machine learning models (CNN & CF) are used for diagnosing emotions to recommend mental health treatment plans

Chen-Yu Hsia & Wanrou Yang

Purpose

Mental-health diagnosis often relies on subjective clinician judgment, patient self-reporting, and observational assessments. These methods can be time-consuming and inconsistent across providers.
Machine learning offers new opportunities for objective analysis:

- CNNs can extract emotional information from facial expressions.
- Collaborative Filtering (CF) can identify therapies that worked for patients with similar symptom patterns.

Goal: Build a prototype that (1) recognizes emotion patterns using CNNs and (2) recommends personalized therapy options using a hybrid CF model.

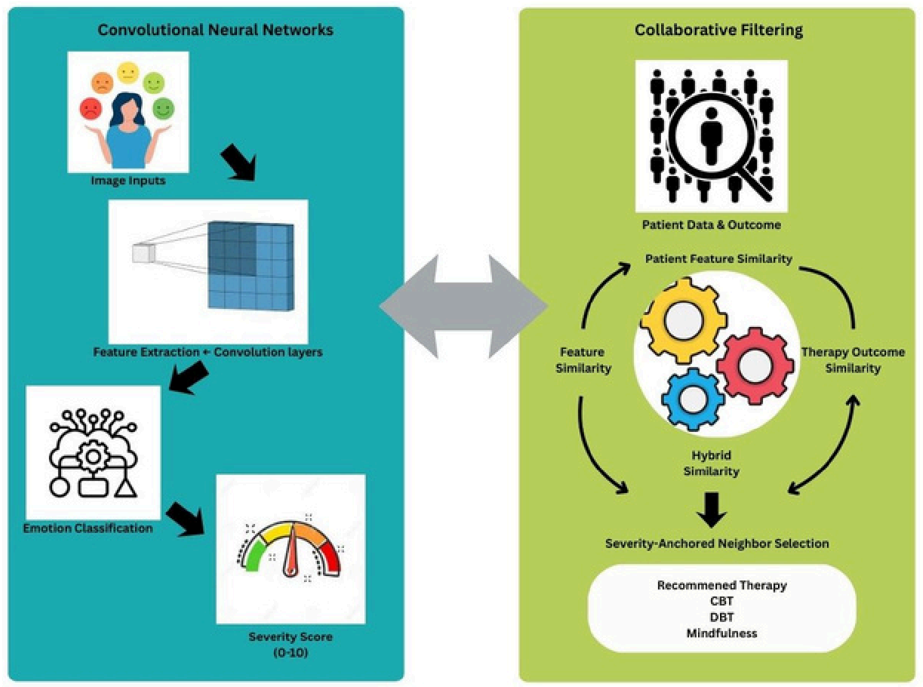
Methods

CNN for Emotion Classification

- Preprocessing: grayscale, resize to 128×128, normalization
- Data augmentation: rotation, flip, zoom
- Architecture:
  - Conv16 + MaxPool
  - Conv32 + MaxPool
  - Dense 128 → Dense 64 → Softmax (7 classes)
- Final Accuracy: ~28% (≈2× random chance), but unstable on real-world images.

Collaborative Filtering for Therapy Recommendation
Hybrid CF = 0.6 × feature similarity + 0.4 × therapy-outcome similarity
Steps:

- Clean and encode outcomes (Improved = 1, No Change = 0, Deteriorated = –1)
- Construct patient × therapy rating matrix
- Compute cosine similarity between patients and between therapy-outcome patterns
- Hybrid similarity = weighted combination
- Anchor on closest severity level
- Pick top-8 most similar patients
- Recommend therapy with highest weighted improvement score



Data

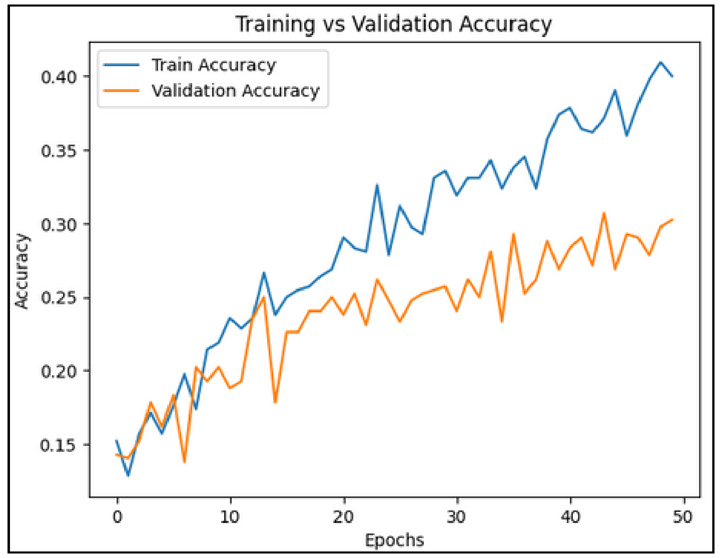
CNN Dataset

- Source: Kaggle Facial Expression Dataset (https://www.kaggle.com/datasets/msambare/fer2013)
- 7 labels: angry, disgust, fear, happy, neutral, sad, surprise
- 420 processed images (60/class)

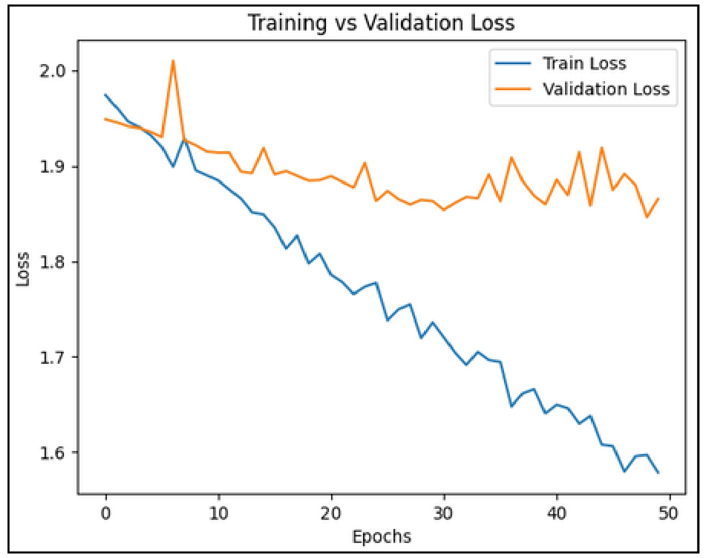
CF Dataset

- Source: Kaggle Mental Health Treatment Dataset (https://www.kaggle.com/datasets/uom190346a/mental-health-diagnosis-and-treatment-monitoring)
- Includes:
  - Severity (1–10), Therapy Type, Outcome
  - Mood scores, stress, sleep, activity, adherence, age
- Outcome encoded as:
  - Improved = 1
  - No Change = 0
  - Deteriorated = −1

Results



The graph shows both the trend of training accuracy steadily improving from 14% to 32% and the validation accuracy also improving, reaching 28%. However, the validation accuracy did not improve as much as the training accuracy due to the small test set size. Both accuracies show clear improvement, indicating that the CNN is learning distinguishable features for each emotion category. When we compare with random accuracy 1/7, it is about 14% and our model reaches 28% validation accuracy, nearly double random, proving it learned more than chance. Extending training to 50 epochs and using augmentation improved performance: training accuracy reached its highest (32%), and validation accuracy reached its highest (28%).



The graph shows that the training loss decreases continuously from ~1.97 to ~1.58 over 50 epochs. Validation loss decreases initially, then plateaus around 1.88–1.92, and fluctuates slightly; moreover, a noticeable gap forms between the training and validation losses after ~20 epochs. The negative slope of both the training and validation losses throughout all 50 epochs indicates that the model continues to minimize error on the training samples. The validation loss drops from ~1.95 to ~1.88, but does not continue to decrease. This means the model improves generalization early but can no longer reduce validation error after about epochs 20–25. The model successfully learns from the dataset but begins to overfit due to limited data volume. Despite this, the validation loss stabilizing (not exploding) is a positive sign that the model still generalizes to some extent.

Background

Mental-health diagnosis often depends on subjective self-reports and clinician observation, which can be inconsistent and difficult to standardize. Advances in machine learning provide new opportunities to analyze emotional and behavioral signals objectively. Convolutional Neural Networks (CNNs) have been widely used in affective computing and clinical imaging because they can automatically learn facial and spatial patterns linked to emotional states. Prior research shows that CNN-based systems can detect cues related to depression, anxiety, ADHD, and other conditions. Collaborative Filtering (CF), commonly used for recommendations, has recently been applied to healthcare to personalize treatment decisions. By identifying patterns across individuals with similar symptoms or outcomes, CF models can help match patients with effective therapy options. Our project builds on these ideas by using a CNN to classify facial expressions and exploring CF as a way to recommend personalized mental-health interventions.

Conclusion

- CNNs can support mental-health assessment, but require large, diverse datasets to generalize properly.
- Hybrid CF can provide personalized therapy suggestions that align with patient similarity patterns and historical outcomes.
- Combining emotion detection + therapy prediction forms a prototype framework for ML-assisted clinical decision support.

Future Work

- Expand the CNN training dataset with more diverse, real-world facial-expression images
- Explore deeper or pre-trained architectures to improve emotion classification
- Enhance preprocessing and data augmentation to reduce overfitting and improve generalization
- Convert CNN outputs into a stable, interpretable severity score that directly feeds the CF model
- Use larger and more representative clinical datasets for CF to reduce sparsity and bias
- Incorporate temporal patient information to capture progress over time
- Add uncertainty or confidence scores to therapy recommendations

References

Mazlan, Idayati, Noraswaliza Abdullah, and Norashikin Ahmad. "Exploring the impact of hybrid recommender systems on personalized mental health recommendations." International Journal of Advanced Computer Science and Applications 14, no. 6 (January 1, 2023). https://doi.org/10.14569/ijacsa.2023.0140699.

P. Chinnasamy et al., "Health Recommendation System Using Deep Learning-based Collaborative Filtering," Heliyon 9, no. 12 (November 24, 2023): e22844, https://doi.org/10.1016/j.heliyon.2023.e22844.

Su, Chengcheng, Zhiguo Xu, Jyotishman Pathak, et al. "Deep Learning in Mental Health Outcome Research: A Scoping Review." Translational Psychiatry 10, no. 116 (2020). https://doi.org/10.1038/s41398-020-0780-3.

Geng, Xiang-Fei, and Junhai Xu. "Application of Autoencoder in Depression Diagnosis." DEStech Transactions on Computer Science and Engineering (2017). https://doi.org/10.12783/dtcse/csma2017/17335.