Project Report: Mental Health Self-Analysis Model

Executive Summary

This project implements a machine learning-based mental health analysis tool that can predict potential mental health conditions based on user input. The system combines BERT-based natural language processing with traditional machine learning techniques to provide accurate classifications and actionable recommendations.

Key Achievements:

- Developed a production-ready model with 96% accuracy.
- Created an intuitive Streamlit-based web interface.
- Implemented real-time analysis and recommendations.
- Ensured responsible handling of mental health data.

Technical Implementation

1. Model Architecture

The system uses a hybrid approach combining:

- BERT (bert-base-uncased) for text feature extraction
- Random Forest Classifier for final prediction
- LDA (Latent Dirichlet Allocation) for topic modeling

Random Forest Performance:

	precision	recall	f1-score
Class 0	1.00	0.98	0.99
Class 1	0.96	0.94	0.95
Class 2	1.00	0.99	0.99
Class 3	0.93	0.97	0.95

Class 4 0.95 0.96 0.96
Accuracy: 0.96

1. Data Processing Pipeline

- Text Preprocessing:
- Lowercase conversion
- Special character removal
- Whitespace normalization
- BERT tokenization

2. Feature Extraction:

- BERT embeddings (768-dimensional vectors)
- Topic modeling using LDA
- Custom vectorization for text analysis

3. Classification Categories

The model identifies five primary mental health patterns:

- Anxiety and Stress Related Symptoms
- Depression and Mood Related Issues
- Sleep and Energy Problems
- Self-esteem and Personal Worth Issues
- Social and Relationship Concerns

User Interface & Deployment

1. Web Interface Features

- Clean, accessible design using Streamlit
- Real-time analysis of user input
- Color-coded results presentation
- Comprehensive recommendations

- Important disclaimers and resources

2. Safety Measures

- Clear disclaimer about non-diagnostic nature
- Emergency contact information
- Professional help recommendations
- Data privacy considerations

Technical Considerations

1. Model Performance

- Fast inference time (<2 seconds)
- Efficient memory usage
- Graceful error handling
- Cached model loading

2. Scalability Features

- Streamlit caching for model loading
- Efficient batch processing
- Modular code structure
- Error recovery mechanisms

3. Future Improvements

- Model Enhancements:
- Fine-tuning on domain-specific data
- Additional language support
- Severity assessment capabilities
- Interface Updates:
- Mobile optimization
- Offline capability
- Multi-language support
- Progress tracking

Deployment Requirements

1. System Requirements

- Python 3.9+
- 8GB RAM minimum
- GPU optional but recommended
- Storage: ~500MB for models

2. Dependencies

- transformers
- torch
- streamlit
- scikit-learn
- numpy
- joblib

3. Installation Steps

- Clone repository
- Install dependencies
- Download pre-trained models
- Run Streamlit server

Ethical Considerations

1. Data Privacy

- No personal data storage
- Anonymous analysis
- Secure processing

2. Responsible AI

- Clear limitations disclosure
- Professional guidance emphasis
- Regular bias monitoring

3. Accessibility

- Screen reader support
- High contrast options
- Clear language use

Conclusion:

The Mental Health Self-Analysis Model successfully combines modern NLP techniques with traditional machine learning to provide accurate, responsible mental health analysis. The system achieves high accuracy while maintaining user privacy and providing clear guidance for seeking professional help when needed.

The project meets all initial objectives and provides a solid foundation for future enhancements. The modular architecture allows for easy updates and improvements as new requirements emerge.