

MENTAL HEALTH PREDICTION

from Reddit & Twitter Dataset Using
Machine Learning and Deep Learning

BY SUBARNA PAUL

ROLL NO.: 11561723008

Under Supervision of Mr. Subhasis Mallick



**B. P. PODDAR INSTITUTE OF MANAGEMENT & TECHNOLOGY Approved by
AICTE, New Delhi and Affiliated to MAKAUT, W.B.**

Problem Statement

“Individuals With Depression Have Specific Texting Patterns in Social Media.”

How We Can Leverage The Categorical/Descriptive Data Into Quantitative Data To Predict Mental Health Issues?



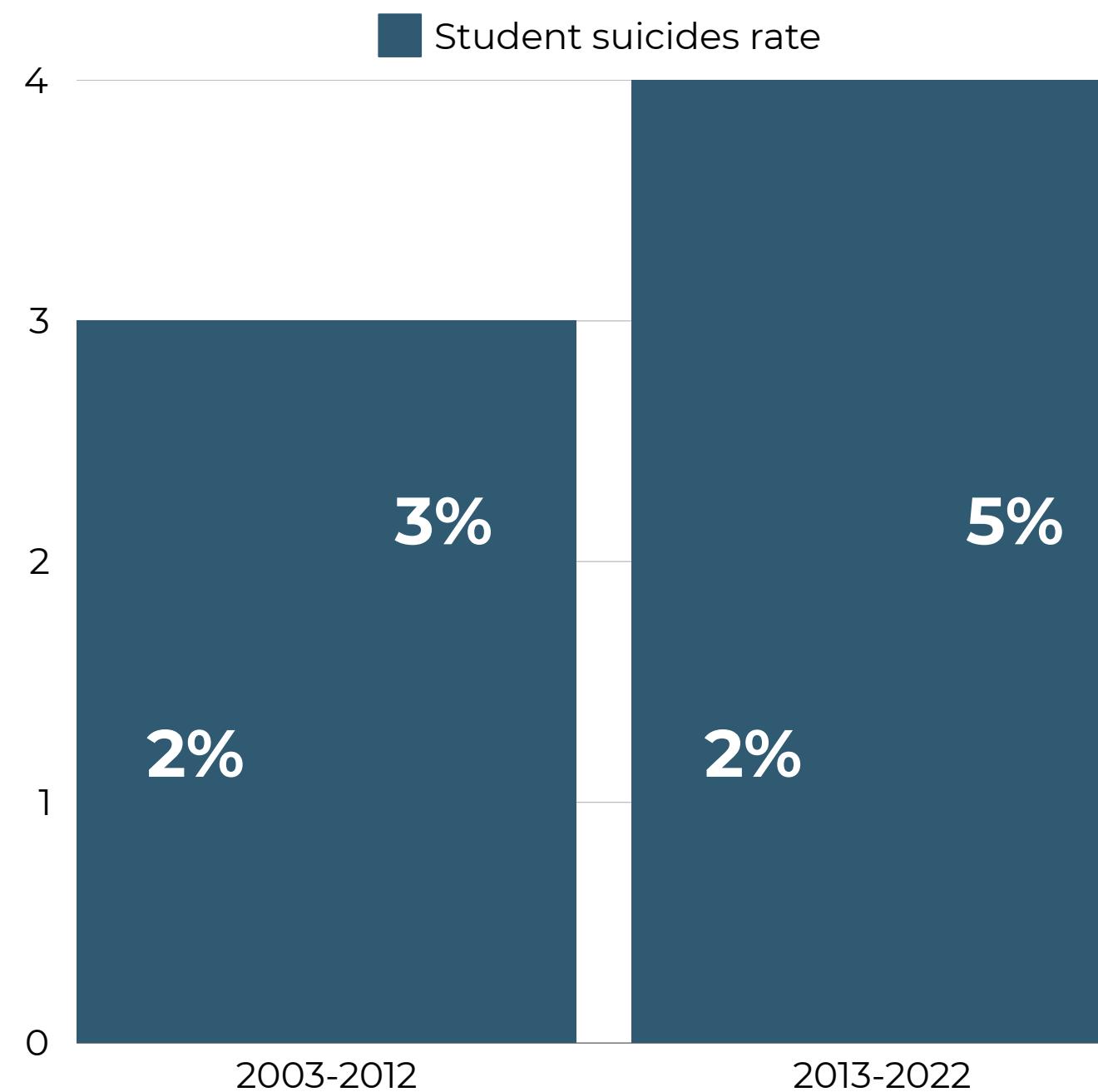
...



OBJECTIVE

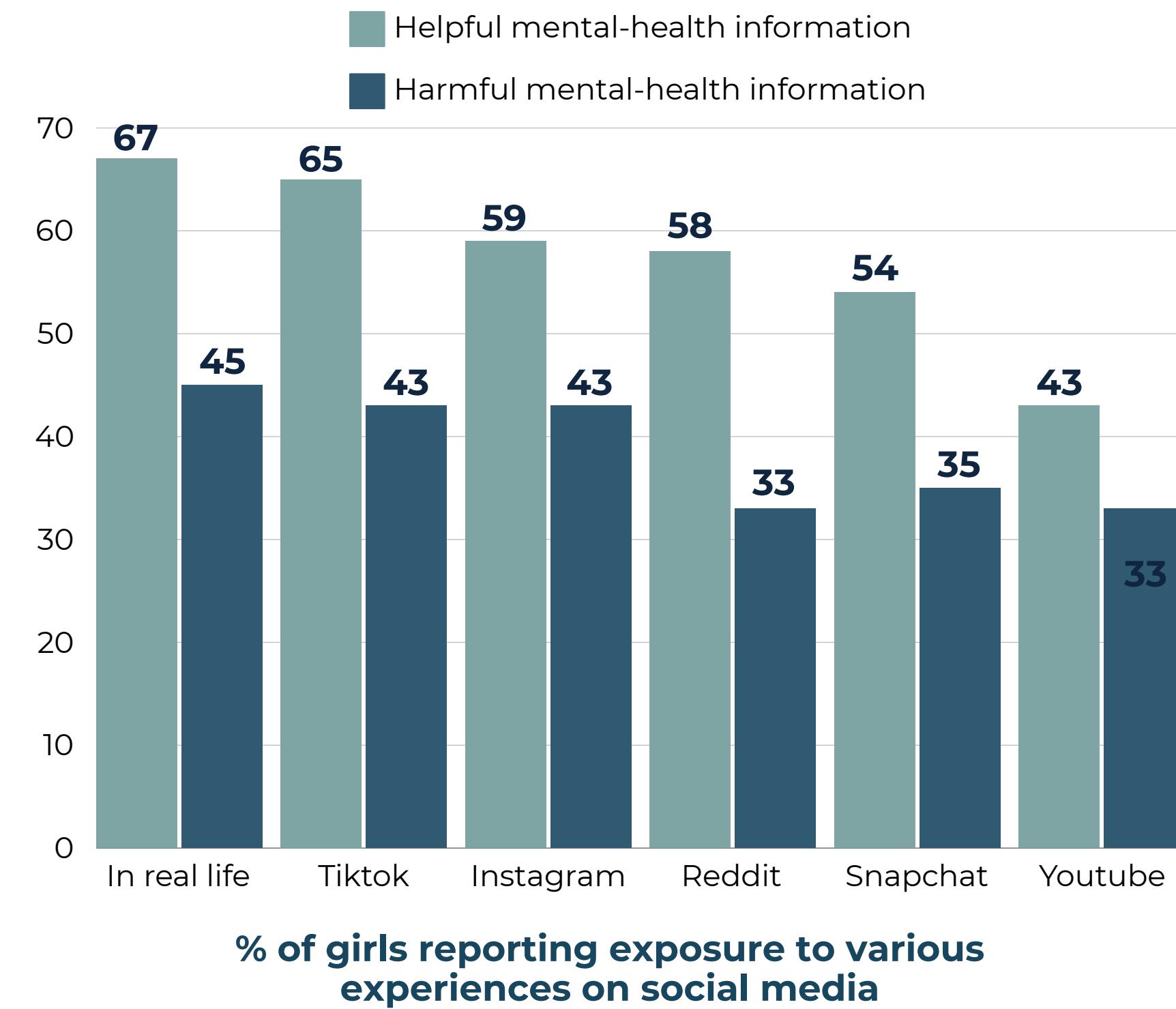
- Dataset Identification (Twitter/Reddit)
- Data Preprocessing
- Dataset Combination (Depression, Anxiety & Normal)
- Model Selection and Training (Machine Learning & Deep Learning)
- Proper Conclusion

Compound Average Growth Rate of Student Suicides



Overall suicides and population (0-24 year old)

Social media usage intersects with mental issues and suicide



MOTIVATION

YEAR	AUTHORS	DATASETS	SAMPLE SIZE	ALGORITHM	METRICS
2020	JinaKim, Jieon Lee, Eunil Park, J Han	Reddit	4,88,472	XgBoost-CNN	Accuracy: 89%, F1-Score: 79.49%
2021	Almars et al.	Twitter	6,000	Bi-directional LSTM	83% Overall Accuracy
2023	Ma et al.	WeChat, Weibo, phone recording	10400 tweets	LSTM, BERT	91.59%, improving the model's stability

MOTIVATION

YEAR	AUTHORS	DATASETS	SAMPLE SIZE	ALGORITHM	METRICS
2023	Wani et al.	Facebook, Twitter, YouTube	52,944	Word2Vec LSTM and Word2Vec CNN	99.02% & 99.01% accuracy respectively
2023	Ghosh et al.	Facebook, Twitter, YouTube	15,031	LSTM-CNN	92.63% sensitivity & 95.12% specificity
2024	Masab A. Mansoor, Kashif Ansari	Twitter, Reddit, Facebook,	996,452 social media posts	multi-modal deep learning model	Accuracy 89.3%, F1 scores 0.827

...

METHODS USED



**Machine Learning -
SVM, Random Forest,
XGBoost**

**Deep Learning -
RNN, LSTM**

DATA OVERVIEW: REDDIT

	Title	Text	Subreddit
0	exposure does not work!	I have struggled with social anxiety from chil...	Anxiety
1	Panic attack? derealization? can't go to docto...	Back in March (I know, a while ago D:), I woke...	Anxiety
2	How long can a panic attack last?!	I've been withdrawing from medicines lately (e...	Anxiety
3	Stepping stones	First time poster, long time lurker. \n\nI've ...	Anxiety
4	Coping with anxiety over climate change, on th...	Hi all,\n\nI made a throwaway account as my ma...	Anxiety
...
488467	Religion and paranoid conversation	Ok! I've got a rant to make and a conversation...	schizophrenia
488468	Isn't this a little weird for being tardive dy...	When I wake up, I rarely have any tics in my t...	schizophrenia
488469	Is this a symptom of schizophrenia or somethin...	I'm not sure if this is schizophrenia (or perh...	schizophrenia
488470	Anyone seen the new Black Mirror episode Bande...	Although I don't think it was stated what Stef...	schizophrenia
488471	My girlfriend of over 1 year and I just broke up	It's really tough right now. We broke up due t...	schizophrenia

488472 rows × 3 columns

DATA OVERVIEW : TWITTER

	id	1	2	twits
0	2401	Borderlands	Positive	I am coming to the borders and I will kill you...
1	2401	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	2401	Borderlands	Positive	im coming on borderlands and i will murder you...
3	2401	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	2401	Borderlands	Positive	im getting into borderlands and i can murder y...
...
74676	9200	Nvidia	Positive	Just realized that the Windows partition of my...
74677	9200	Nvidia	Positive	Just realized that my Mac window partition is ...
74678	9200	Nvidia	Positive	Just realized the windows partition of my Mac ...
74679	9200	Nvidia	Positive	Just realized between the windows partition of...
74680	9200	Nvidia	Positive	Just like the windows partition of my Mac is l...

74681 rows × 4 columns

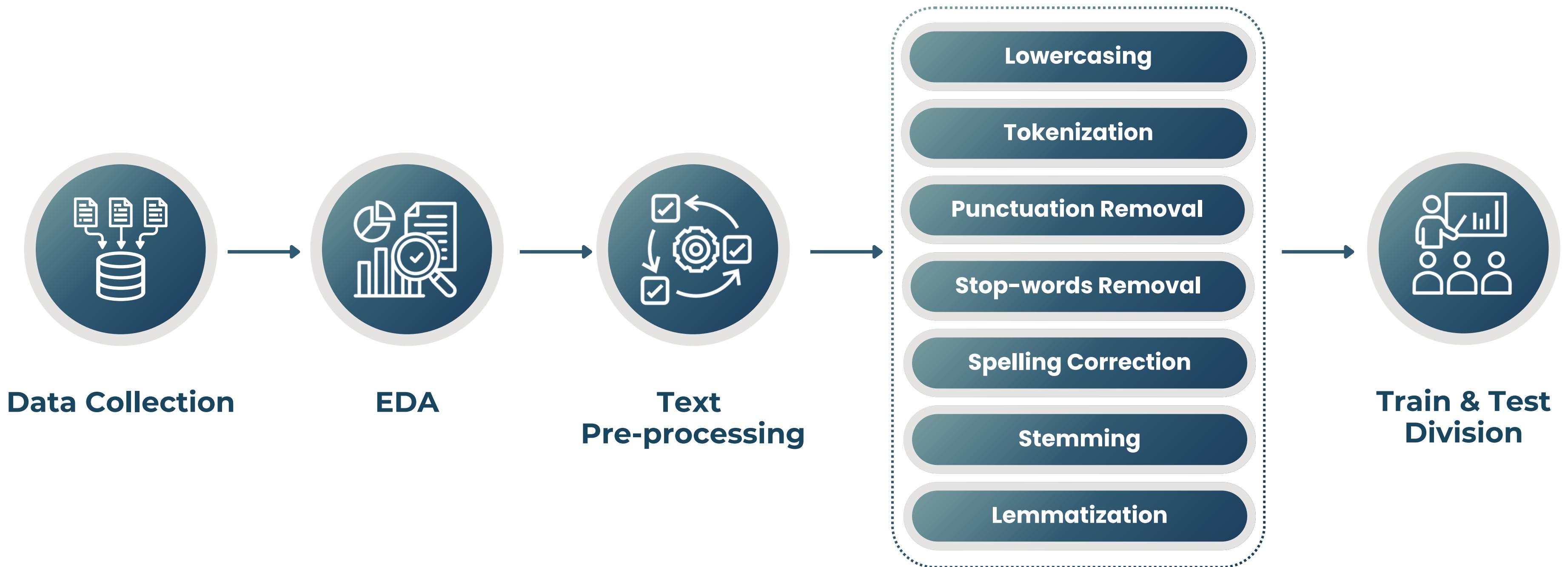
DATA PROCESSING

	Cleaned_Text	class
0	holy crap guy last day anyone else little sad ...	anxiety
1	wrong major first three year college racked lo...	anxiety
2	today spoken word assessment english go well i...	anxiety
3	2017 lmao wtf	normal
4	im 27 depressed since high school though past ...	depression
...
35995	strange title sorry wanted know anyone else fe...	anxiety
35996	hi first ever post reddit took like 5 minute f...	depression
35997	one ever really talk make life worth living gr...	depression
35998	im sorry isnt appropriate place really need le...	depression
35999	three month ago got medically discharged air f...	anxiety

36000 rows × 2 columns

- Classes: Depression, Anxiety, Normal
- Data Points: 36,000
- Depression: 12,000
- Anxiety: 12,000
- Normal: 12,000

FLOW OF WORK TILL NOW



FLOW OF WORK TILL NOW



RESULT (Multiclass Classification)

...

Machine Learning Models:

Accuracy: 0.8861

Classification Report:

	precision	recall	f1-score	support
anxiety	0.88	0.83	0.85	3588
depression	0.83	0.86	0.84	3603
normal	0.95	0.97	0.96	3491

Accuracy: 0.8736

Classification Report:

	precision	recall	f1-score	support
anxiety	0.88	0.81	0.84	3588
depression	0.83	0.83	0.83	3603
normal	0.91	0.98	0.94	3491

Support Vector Machine Model (88.61%)

Accuracy: 0.8554

Classification Report:

	precision	recall	f1-score	support
anxiety	0.86	0.79	0.82	3588
depression	0.82	0.79	0.80	3603
normal	0.89	0.99	0.94	3491

XGBoost Model (87.36%)

Random Forest Model (85.54%)

RESULT

Deep Learning Models:

Accuracy: 0.8609

Classification Report:

	precision	recall	f1-score	support
0	0.82	0.83	0.83	3588
1	0.82	0.80	0.81	3603
2	0.94	0.96	0.95	3491

RNN Model

ACCURACY - 86.09%

Accuracy: 0.8738

Classification Report:

	precision	recall	f1-score	support
0	0.86	0.84	0.85	3588
1	0.84	0.81	0.82	3603
2	0.92	0.98	0.95	3491

LSTM Model

ACCURACY - 87.38%

...

RESULT (Binary Classification)

...

Subreddit	Class	XGBoost Model		LSTM Model	
		F1 Score	Accuracy	F1 Score	Accuracy
r/depression	Depression	96	95.69%	95	95.26%
	Normal	96		95	
r/Anxiety	Anxiety	97	97.37%	95	94.86%
	Normal	97		94	
r/bipolar	Bipolar	96	95.89%	96	96.02%
	Normal	96		96	
r/BPD	BPD	96	96.17%	97	97.05%
	Normal	96		97	
r/schizophrenia	Schizophrenia	94	94.05%	95	95.56%
	Normal	94		96	
r/autism	Autism	96	95.66%	95	94.96%
	Normal	96		95	

LIMITATIONS

Key Challenges in Machine Learning in Healthcare:

- **Dependence on Data:** Requires large, well-annotated datasets, often difficult to obtain, especially sensitive health data.
- **Misclassification:** Misdiagnosis by getting negative words as positive way or the vice versa.
- **Generalization Issue:** Models may excel on training data but struggle with new data due to differing diagnostic methods.
- **Fairness and Bias:** Bias in data can lead to uneven performance; diverse data is essential to prevent discrimination.



FUTURE PLAN

- Development and Improvement of LLM for Mental Health Analysis
- Enhancing Symptom Analysis and Prediction
- Generating Remedies and Supportive Interventions
- Application Implementation on Social Media Platforms
- Result Optimization and Continuous Improvement
- Collaboration with Healthcare Providers



REFERENCE

- Depression and anxiety detection from blog posts data. By Yevhen Tyshchenko. (2018)
- Sentiment Analysis of Social Media Contents using Machine Learning Algorithms. By R.I. Ahmed, M.J. Iqbal, and M. Bakhsh. (2020)
- A deep learning model for detecting mental illness from user content on social media. By Jina Kim, Jieon Lee, Eunil Park & Jinyoung Han. (2020)
- Automatic Detection and Classification of Mental Illnesses from General Social Media Texts. By Anca Dinu and Andreea-Codrina Moldovan. (2021)
- Machine learning of language use on Twitter reveals weak and non-specific predictions. By Sean W. Kelley, Caoimhe Ní Mhaonaigh, Louise Burke, Robert Whelan and Claire M. Gillan. (2022)

REFERENCE

- Web references:
 - <https://www.yalemedicine.org/news/social-media-teen-mental-health-a-parents-guide>
 - <https://www.nature.com/articles/s41598-020-68764-y>
 - <https://www.igi-global.com/article/machine-learning-for-depression-detection-on-web-and-social-media/342126>
 - <https://www.geeksforgeeks.org/machine-learning/>
 - <https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/>
 - <https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/>

THANK YOU
