CIIS 2025
Advancing Investigation in the Digital Age

National
CyberShield Hackathon 2025

X-AntiIndia Detector

- Problem Statement - Develop a system to detect anti-India campaigns on digital platforms by building a keyword database and using engagement analysis to identify key drivers, supporting early detections.
- Team Name - ***lyf will be harder***
- College Name - SVKM's NMIMS Indore
- Idea: Multi-Model AI, leveraging NLP and Computer Vision, for real-time stream analysis of video and image-based content. Generate predictive intelligence to classify disinformation and malicious narratives.

Idea

- Proposed Solution : A system to detect anti-India campaigns using a dynamic keyword database, Selenium-scraped training data, X API for testing data, NLP classification, engagement analysis, YouTube transcript analysis, and a visualization dashboard.

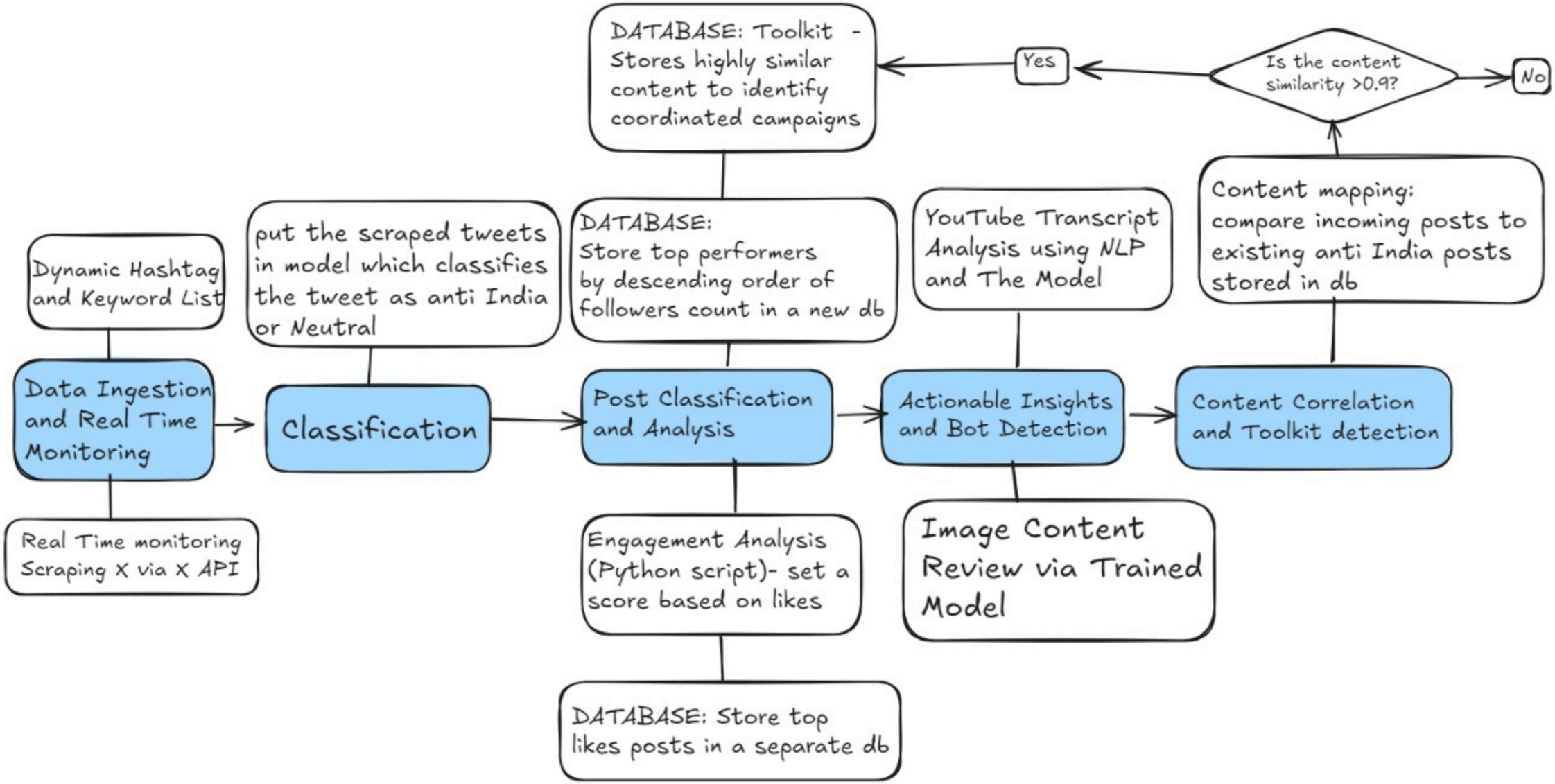
- Detailed explanation of the proposed solution: Build a dynamic keyword list; Scrape training data with Selenium (500 posts/keyword); Use X API for real-time testing posts; Classify on an anti-India scale with distilbert-base-uncased; Analyze engagement (likes, retweets); Extract YouTube transcripts; Visualize in a Streamlit dashboard.

- Innovation and uniqueness of the solution: Dynamic keyword updates from scraped data; Multi-platform (X, YouTube); Image analysis with Gemini; Ethical logging; Shifted from X API to Selenium for training data due to cost.

Technical Approach

- Technologies to be used:
  - *Programming languages: Python 3.12 (Pandas, Transformers, Selenium)*
  - *Frameworks: Hugging Face(distilbert-base-uncased), pytube, youtube-transcript-api, google-generativeai*
  - *Database: SQLite*
  - *Dashboard: Streamlit*
  - *Other: Weights and Biases*

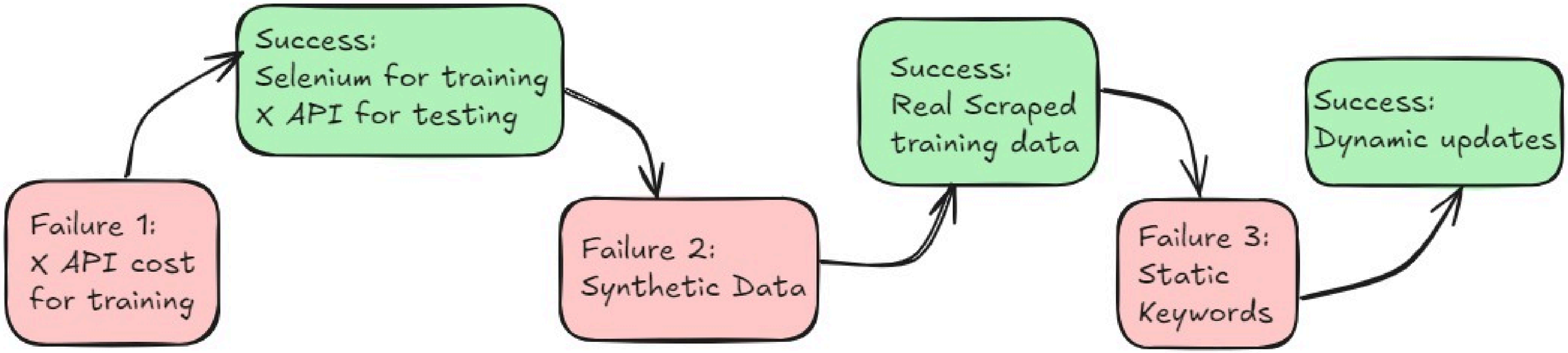
Methodology and Process for Implementation



Feasibility And Viability

- Analysis of the feasibility of the idea: Feasible with open-source tools (Selenium, Hugging Face); Built within hackathon timeframe using 500 training posts/keyword and X API testing data; Supports multi-lang via xlm-roberta.
- Potential challenges and risks: X API costs for training data led to Selenium switch (ban risk); Initial X API failure for training; Model accuracy on nuanced content.
- Strategies for overcoming these challenges: Use efficient models like *DistilBERT* to reduce compute needs; Implement error handling for API limits; Fine-tune on India-specific datasets; Validate with manual checks and iterative testing. Use Selenium with delays/proxies for training; Switched to X API for testing; Fine-tune with augmentation; Manual validation.

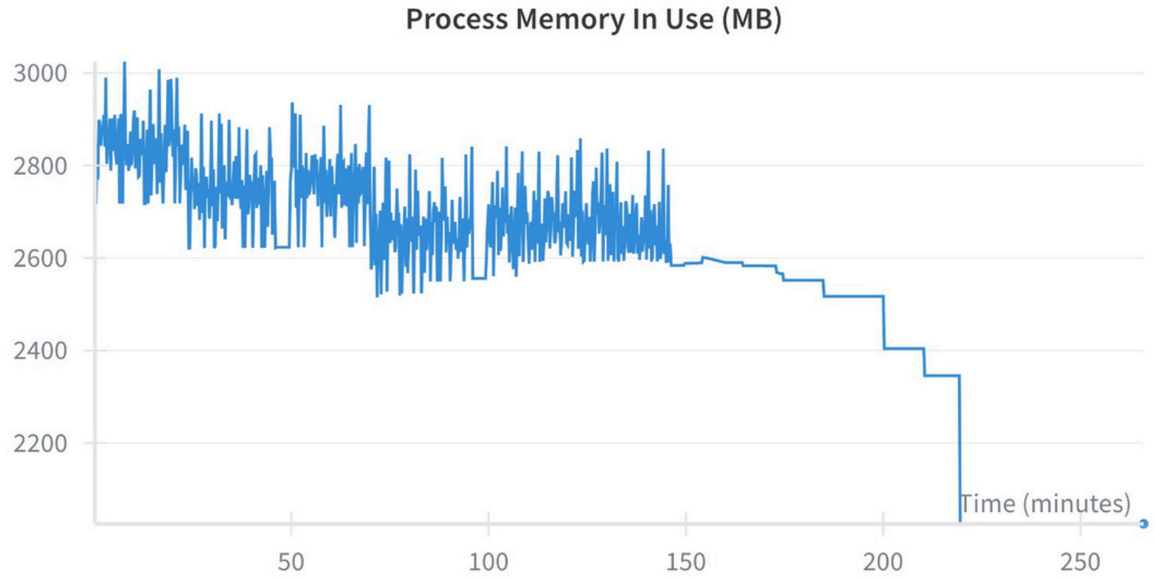
Failures to Success

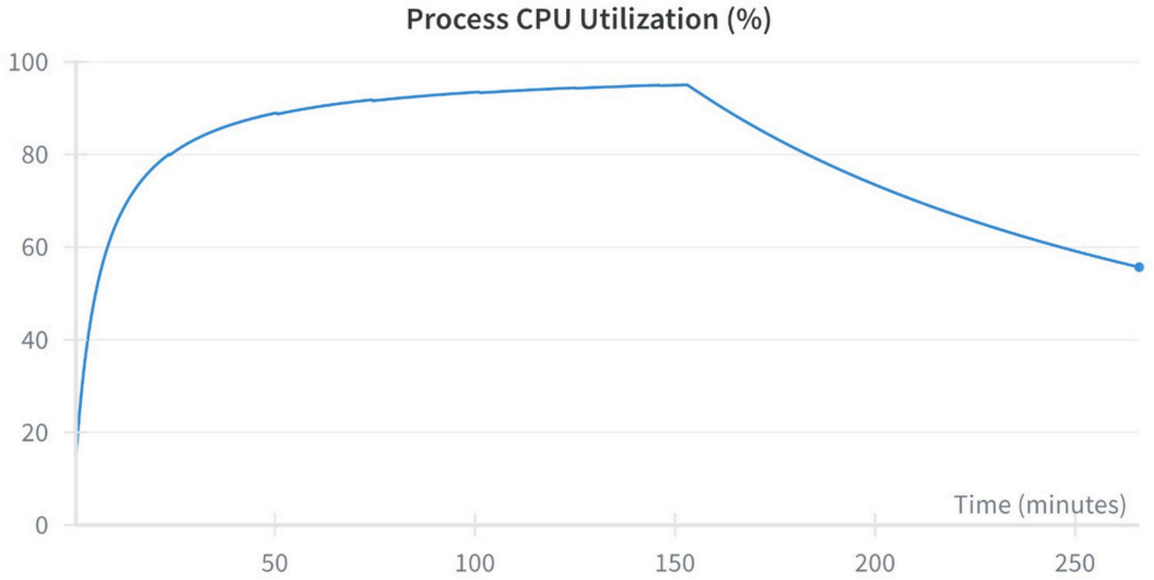


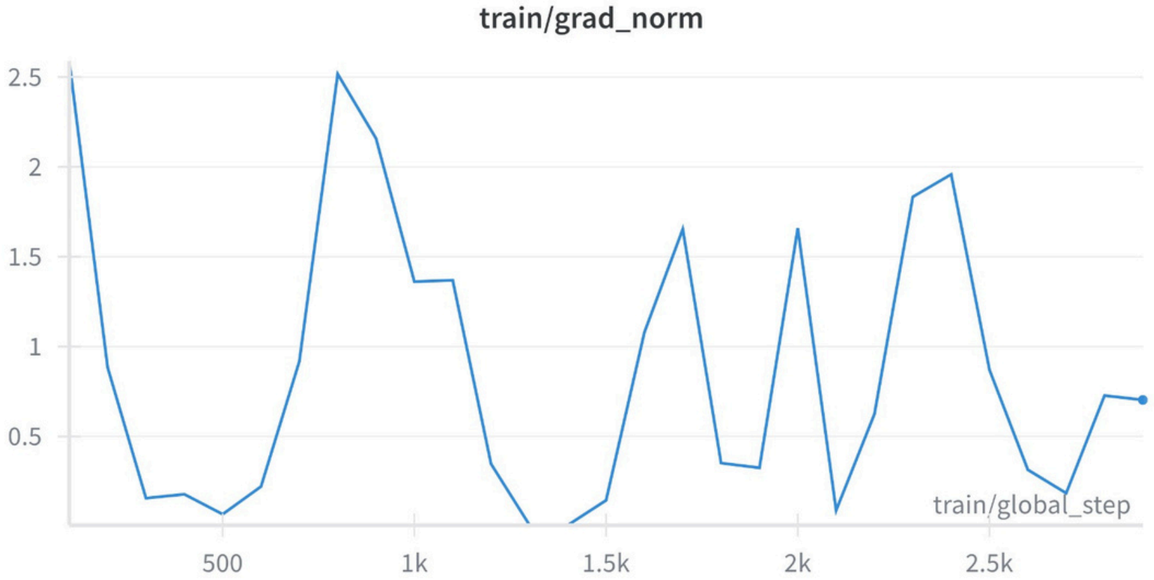
System & Training Metrics

National CyberShield Hackathon 2025

CIIS 2025 Advancing Investigation in the Digital Age







Impact and Benefits

- Potential impact on the target audience: Empowers security agencies, researchers, and social media users to identify misinformation networks; Raises awareness about bot-amplified propaganda.

- Benefits of the solution (social, economic, environmental, etc.):
  - Social - *Reduces spread of anti-national narratives.*
  - Economic - *Aids in efficient resource allocation for cyber investigations.*
  - Environmental - *Minimal impact, as it's software-based with low energy models like DistilBERT.*

Future Endeavours

1. Content Classification (Law Violation): Detect hate/anti-India posts using NLP (keywords + similarity).

Map flagged posts by Indian laws:

IPC 124A → Sedition

IPC 153A → Promoting enmity

IPC 295A → Religious hate speech

IT Act 66A/67 → Offensive digital content

2. IP & Bot Farm Detector : Track IP addresses of Users/Bots.

- *If >10 flagged posts from same IP* → Bot Farming Alert
- Fast repetitive posting → Automated behavior suspicion.

Model Output Status

National CyberShield Hackathon 2025

CIIS 2025 Advancing Investigation in the Digital Age

```
[23] sample_tweet_anti = "The 'Sanchi Stupa' in Madhya Pradesh is a great Buddhist complex, famous for its Great Stupa."
     prediction_anti = predict_tweet_sentiment(sample_tweet_anti)
     print(f'Tweet: "{sample_tweet_anti}"\nPrediction: {prediction_anti}\n')

     Tweet: "The 'Sanchi Stupa' in Madhya Pradesh is a great Buddhist complex, famous for its Great Stupa."
     Prediction: Neutral


[24] sample_tweet_anti = "I see Indians as disruptors because they invading. Remove them all"
     prediction_anti = predict_tweet_sentiment(sample_tweet_anti)
     print(f'Tweet: "{sample_tweet_anti}"\nPrediction: {prediction_anti}\n')

     Tweet: "I see Indians as disruptors because they invading. Remove them all"
     Prediction: Anti-Indian
```

The model shows strong performance with decreasing loss and perfect accuracy, over three epochs. It accurately predicts sentiments, labeling a neutral tweet as "Neutral" and a hostile one as "Anti-Indian."

Epoch	Training Loss	Validation Loss	Accuracy
1	0.105100	0.120684	0.956167
2	0.101700	0.127754	0.956167
3	0.105500	0.118210	0.956167

Research and References

- *[1] X API Documentation. Available at: https://developer.x.com/en/docs/twitter-api. Accessed: 31 August 2025.*
- *[2] Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108. Accessed: 31 August 2025.*
- *[3] Hugging Face Transformers library. Available at: https://huggingface.co/transformers. Accessed: 31 August 2025.*
- *[4] Scikit-learn for clustering : Available at: https://scikit-learn.org/stable/modules/clustering.html. Accessed: 31 August 2025.*
- *[5] Plotly for visualizations. Available at: https://plotly.com/python/. Accessed: 31 August 2025.*

Team Member Details

Reg.No/ Enrol.No	Name	Year/ Semester	Name of College/ Department	City/ State	Male/ Female	Mail ID	Mobile No
1.	**Anupam Kanoongo**	II/III	SVKM's NMIMS Indore | STME | Computer Engineering	Indore (Madhya Pradesh)	M	anupam.kanoongo311@nmims.in	8889116054
2.	**Priyansh Saxena**	II/III	SVKM's NMIMS Indore | STME | Computer Engineering	Indore (Madhya Pradesh)	M	priyansh.saxena807@nmims.in	7566092515
3.	**Srishti Jain**	II/III	SVKM's NMIMS Indore | STME | Computer Engineering	Indore (Madhya Pradesh)	F	srishti.jain748@nmims.in	9294879608
4.	**Jayendra Singh Rawat**	II/III	SVKM's NMIMS Indore | STME | Computer Engineering	Indore (Madhya Pradesh)	M	jayendra.singhrawat263@nmims.in	9131037755
5.	**Tanay Trivedi**	II/III	SVKM's NMIMS Indore | STME | Computer Engineering	Indore (Madhya Pradesh)	M	tanay.trivedi373@nmims.in	8643082119

Mentor Detail

Staff Id	Name	School/ Department	Mail ID	Phone . No
1.	**Dr. Vikas Kumar Jain**	School Of Technology Management & Engineering	vikas.jain@nmims.edu	9981008680

Thank You!