Applicant's Name	NAYUNI INDU ESWAR SHIVANI	App. No.	*1
Field of Study *Mark one field.	**CS SY CN IT SE PM**		
Prospective Research Adviser	Name	Seal or Signature *2	
	RAGE UDAY KIRAN		
	Recommended Conversion Courses *3		
	CV1 CV2 CV3 CV4 CV5 CV6 CV7 CV8		

Research Plan

Write in English within two pages. The entry fields can be expanded.

Title/Research Subject:

Integrating Air Pollution Knowledge Graphs with LLMs for Advanced Air Quality Insights

Background/Related Works:

Air pollution monitoring traditionally relied on sensor networks and statistical models, which face challenges in scalability and real-time data processing. Recent advancements in **AI**, particularly Large Language Models (LLMs) like GPT-3 and Gemini, have shown potential in analysing vast amounts of unstructured data, extracting insights, and making predictions from environmental datasets. These models, when combined with Knowledge Graphs (KGs), which structure relationships between pollution sources, weather patterns, and other environmental factors, enable the creation of intelligent systems for better decision-making and air pollution management.As research in this area continues to evolve, it holds the promise of transforming air quality monitoring systems, facilitating more effective and targeted interventions for mitigating the impacts of air pollution on public health and the environment.

Objectives/Expected Contributions:

1. **Knowledge Extraction**: Use LLMs like Gemini and GPT-3 to extract insights from environmental data, enhancing the understanding of air pollution trends and causes.
2. **Real-Time Data Integration**: Combine real-time pollution and weather data to improve prediction accuracy.
3. **Predictive Modeling**: Develop AI models to predict future air pollution levels for proactive management.
4. **Knowledge Graph Construction**: Build a knowledge graph linking pollution sources, weather, and other factors to improve decision-making.
5. **Scalability**: Apply cloud computing to scale the system for real-time monitoring in large areas.
6. **Advanced Query System**: Develop a system for stakeholders to interact with the knowledge graph and access data-driven insights.

Approach/Methodology:

1. LLM Integration for Knowledge Extraction

- Use pre-trained LLMs (e.g., GPT-3, Gemini) to extract insights from environmental data and answer specific queries related to air pollution and its contributing factors.

2. Predictive Analytics

- Develop predictive models to forecast air quality based on historical trends and real-time data.
- Apply machine learning techniques like regression or time-series forecasting for accurate predictions.

3. Query System Development

- Implement an interactive query system for stakeholders to retrieve pollution data, trends, and predictive insights.

4. Real-Time Data Integration

- Integrate real-time environmental data with predictive models for continuous monitoring and real-time air quality forecasting.

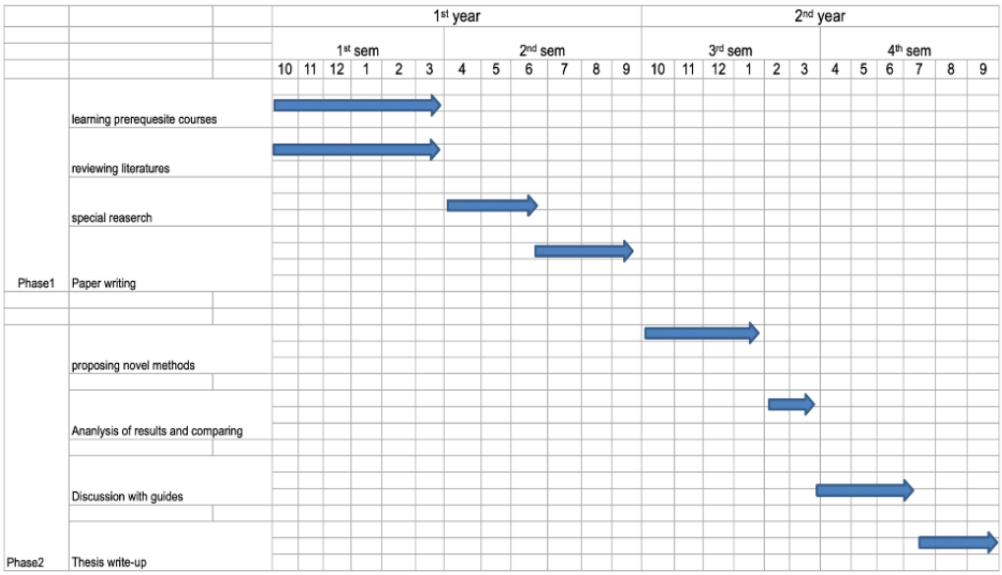
5. Validation

- Validate predictive models using performance metrics such as RMSE, MAE, and accuracy to ensure their reliability and precision in real-world applications.
- Conduct cross-validation on historical data to assess model robustness and generalization.

6. Cloud-Based Deployment

- Deploy the system on scalable cloud infrastructure to process large volumes of data, enabling real-time analysis and broad accessibility for users.

Timeline:



*1 Leave the "App. No." field blank.

*2 *3 Prior to making an application, **it is necessary to personally obtain approval from your prospective research adviser** (and recommendations for conversion courses as required) regarding the research plan.
If you are in a remote place and not able to obtain a seal or a signature, a copy of an email that verifies their agreement should be attached.

Conversion Courses

CV1	Logic Circuit Design	CV5	Algorithms and Data Structures
CV2	Programming Languages	CV6	Formal Languages and Compilers
CV3	Operating Systems	CV7	Database Management Systems
CV4	Computer Architecture	CV8	Computer Graphics