**22AIE315:Natural Language Processing**

**Guidelines of Term Project**

1. **Decide your team**
   1. Decide your team members .No: of Members in a team : 2 to 3 (maximum)
   2. Give a name to your team.
2. **Select a domain of the problem** : Healthcare/Education/Software engineering/Material Science/Finance/Law/Tourism/industrial automation/ Bioinformatics/ Cybersecurity & Threat Intelligence /Astronomy & Space Science (LLM for space data interpretation)/Climate Science & Sustainability (AI-driven environmental modeling), Pharmaceuticals & Drug Discovery (LLM-assisted molecule design)/Smart Cities & IoT (AI for urban planning and automation), Defense & Military Applications (AI for security analysis),/Retail & E-Commerce (LLMs for personalized shopping experiences)/Gaming & Interactive AI (Agentic LLMs for dynamic gaming )/Cultural Heritage & Linguistics (AI for language preservation & historical document analysis)/or anyother.....
3. **Identify the functionalities of the project (Refer to sample below)**

Try to bring in novel components in the functionality. Do a review of existing tools and work in the problem. Add some new functionalities which can make the project unique.

A sample functionalities list for your reference.

**Project: AI-powered fitness app using LLMs**

Key Functionalities (sample) –

* AI-powered chatbot for fitness queries
* Personalized health tracking and goal setting
* Personalized Workout Generation - Creates adaptive training plans based on user data and fitness goals.
* Health Prediction/Injury Risk Prediction - AI analyzes workout history, fatigue, posture, and heart rate to predict injury risks.
* Health & Lifestyle Predictions - AI estimates how lifestyle habits impact long-term health outcomes.
* Nutritional Guidance - Generates meal plans and explains diet benefits based on health goals.
* Gamification & AI Rewards - AI-driven badges, leaderboards, and rewards for fitness milestones.
* AI-Powered Workout Reports & Performance Analytics, Dashboard

1. **Methodology (Mandatory components):**

We expect you to develop an end to end LLM application with the below mandatory modules of LLM Pipeline

* **Frontend ( Web based interface or Mobile App)**
* **Select a suitable pret+rained LLM with proper justification why the model**
* **Appropriate Fine tuning technique**
* **Appropriate Prompt Engineering Strategy**
* **RAG**
* **RLHF**

..................................................**Milestones**.......................................................

**Milestone 1: Problem Submission**

* Abstract Submission.
* Problem definition.
* Identify the functionalities of the project. Identify some novel functionalities too.
* What makes your application unique when compared to the existing applications

**Rubrics**

* Uniqueness in functionality, Novelty (3 marks)
* Clarity and presentation (2 marks)

**Milestone 2: Background study document ( Phase I)**

* Detailed review on existing Tools that already do what you are planning to do
* Detailed review of Research papers in the domain, which has contributed to the problem that you are dealing with.
* A table chart of the pretrained models existing
* Selection of the best suitable and feasible pretrained model for your work with justification

(2 review papers summary+ 12 reputed and recent journal+ existing tools), Gap Analysis

**Rubrics**

* Selection of Minimum 15 papers including 2 survey papers –4marks
* Gap Analysis- 4 marks
* Document Prepartion:2 marks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Criteria** | **Excellent (4 Marks)** | **Good (3 Marks)** | **Average (2 Marks)** | **Poor (1 Mark)** |
| **Selection & Review of Papers (4 Marks)** | Includes at least 15 papers, including 2 survey papers. Provides detailed summaries, clearly explaining methodologies, key findings, and their relevance to the research problem. | Summarizes most papers well but lacks depth in explaining methodologies or key findings | summarize papers without discussing methodologies or their relevance in detail. Lacks strong justification or connection to the problem. | Minimal effort in reviewing the papers, with vague or unclear summaries. Lacks proper discussion and critical analysis. |
| **Gap Analysis (4 Marks)** | Clearly identifies and explains key research gaps with strong supporting evidence from the literature. Provides a well-structured discussion on the limitations of existing studies | Identifies some research gaps but lacks in-depth analysis. Some connections to the research problem are unclear | Mentions research gaps but provides only a brief or weak justification. Lacks clarity | Fails to properly identify research gaps. The analysis is vague and lacks supporting references. |
| **Document Structure & Presentation (2 Marks)** | Well-organized and properly formatted with clear citations, and a logical flow. | Mostly well-structured with minor flaws in logical flow | The document has inconsistencies, missing citations, or lacks clarity. | Poorly formatted, difficult to follow, with missing citations |

**Milestone 3: MidReview & Report ( Phase II)**

* Detailed explanation of methodology with a good comprehensive block diagram depicting the entire project
* A snapshot of existing tools and papers
* What is going to make your work unique. What are the novel components of the work (Unique functionalities/novelty in methodology etc)

**Rubrics**

* Quality and Quantity of the work, Uniqueness and relevance of the functionality , Novelty in the methodology, Uniqueness in functionality- 5 marks
* Block Diagram:5 marks
* Presentation and Viva(Individual): 10 marks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Criteria** | **Excellent** | **Good** | **Average** | **Poor** |
| **Novelty in Methodology,**  **Uniqueness in Functionality,**  **Block Diagram**  **(5 Marks)** | creative and unique approach. It stands out from existing methods. Block diagram is well organized and technically accurate. | some new ideas, but parts of it are similar to existing methods with useful features. Some of the details of the block diagram is missing. | mostly based on existing methods, with only small improvements. The functionality is basic. Block diagram is not well organized. | The project is not unique and uses a common approach. It does not improve much on existing solutions. poorly made block diagram which does not clearly show the system. |
| **Presentation & Viva (Individual) (10 Marks)** | The student presents clearly and confidently answers questions | The student presents with a good understanding but a few answers wrong. | The presentation is unclear and gives weak answers. | The presentation is confusing and cannot answer questions properly. |

**Milestone 4 : Final Presentation and Demo (Phase III)**

* Final Presentation
* Demo
* Code Review

**Rubrics**

* Code review, demo, Presentation, Viva, Indvidual contribution
* End-User Usability & Interface – Intuitive design, ease of navigation.
* Relevance of the Application – Addresses a real-world problem effectively.
* WOW Factor & Innovation – Novelty, uniqueness, and cutting-edge LLM techniques.

**Milestone 5: Report Submission**

* Report in a paper format (Top Tier conference or Journal Format)- 10 marks

**Rubrics**

* Title is concise, Precise and informative
* Abstract is well structured summarizing key findings, methodology and contribution effectively
* Clear Background, Strong motivation, well defined problem statement
* Comprehensive, well-cited, critical analysis of past work; identifies gaps & contributions effectively.
* Well structured explanation of the methodology highlighting the novelty/Research contribution of the work.
* High-quality visuals, clear labels, properly formatted figures/tables
* Comprehensive Performance Evaluation, benchmarking efficiency, and demonstrating improvements over existing solutions.
* Well Structured and Clear Writing style

...............................................................................................................................

**Few Research Tips**

1. **LLM New Trends**

You may explore new LLM trends that can be incorporated for extra research contributions (few Tips are below)

* Better Reasoning LLM
* Novelty in RAG Methodology for Enhanced Contextual Retrieval
* Innovative Reinforcement Learning Strategies for Enhancing LLM Capabilities
* Innovative Optimization Techniques to Enhance LLM Performance and Scalability
* Models with enhanced logical reasoning and math abilities
* Neuromorphic AI
* Ensemble LLMs
* Agentic LLMs -AI models with autonomous multi-step planning
* Multimodal LLMs (Beyond Text)-Integrating vision, speech, and sensor data
* Enhancing Large Language Models with Knowledge Graphs
* Memory-Augmented LLMs-Long-term memory storage for continuous learning
* Neuro-Symbolic Hybrid AI-Combining LLMs with formal logic and knowledge graphs
* Privacy-Preserving LLMs-Federated learning & differential privacy-based fine-tuning

1. **Tasks that an LLM can do. (This is not a complete list)**

***Chatbot is not the only task LLMs can do***. It can do much more. Explore more

while fixing your functionality. Few example tasks that LLM can do are below

* **Knowledge Representation & Reasoning**
  + Creation of a Decision Tree from Text Input
  + Knowledge Graph Construction & Integration
  + Causal Reasoning & Event Prediction -Determines cause-effect relationships from textual data.
* **Advanced Text Processing & Understanding**
  + Chatbots
  + LLMs excel in analyzing, transforming, and structuring text-based content.
  + Text Summarization & Compression
  + Question Answering (QA) & Contextual Search
  + Text Rewriting & Paraphrasing
  + Named Entity Recognition (NER) & Information Extraction-Identifies names, locations, organizations, dates, and key entities.
  + Sentiment Analysis & Emotion Detection-
    - Classifies sentiment from customer reviews, social media, or surveys.
    - Example: Detecting public perception of a political speech.
* **Prediction & Analytical Tasks**
  + Predictive Text & Auto-Completion
  + Autocompletes sentences, paragraphs, or code.
  + Example: AI-powered email drafting and coding suggestions.
  + Trend & Market Prediction
    - Forecasts economic trends, stock movements, or climate patterns based on textual data.
  + Predicting crypto market fluctuations from news sentiment.
  + Fake News & Misinformation Detection
    - Verifies claims by cross-referencing trusted sources.
  + Legal Document Analysis & Risk Prediction
    - Evaluates contract clauses, legal risks, and compliance.Example: Detecting loopholes in corporate contracts.
  + Medical Diagnosis & Symptom Prediction
    - Uses medical texts + patient data to suggest potential diagnoses.Example: Predicting heart disease risk from patient history + research papers.
* **Data Analytics and Visualisation**
  + Auto-Generated Charts from Text Prompts-Example: "Create a bar chart showing monthly revenue for 2023.
  + Adaptive Visualization Selection
  + Dashboard & Report Generation-Example: Summarize sales performance and generate an interactive report.
* **Security**
  + Detects phishing attacks, malware threats, and cybersecurity risks.
  + Identifies unusual behavior in banking transactions or healthcare claims.
  + Automated Regulatory Compliance Audits
* **Autonomous AI Agents (Agentic LLMs)**
* Plans and executes multi-step tasks independently (e.g., AutoGPT, BabyAGI).Example: AI-powered personal finance assistant that manages expenses.
* Data Extraction & Structuring from PDFs, Spreadsheets, and Web Pages
* Converts unstructured data into structured formats.
* Personal Assistants
  + Handles emails, calendar scheduling, and reminders.Example: AI-based executive assistant for CEOs.
* **Content Creation (Articles, Blogs, Scripts)**
  + Writes long-form articles, news, and creative fiction.
* **Multimodal LLM Applications**
  + Image & Video Captioning
  + Generates context-aware descriptions for images/videos.
  + Speech-to-Text & Text-to-Speech (TTS) Processing
  + Music & Lyrics Generation
  + Composes melodies and lyrics based on input themes.
* **Software Engineering**
  + Code Generation & Auto-Completion
  + Automated Software Documentation
  + Automated Software Designs
  + Automated test case generation
  + Automated Software Testing

And many more..........................

**XXXXXXXXXXXXXXXXX Happy Coding ! All the Best XXXXXXXXXXXXXXXXXXXXX**