StudyBuddyy
Your Personalized AI Learning Companion

Team members

- **Abinash Yadav [THA081BEI002]**
- **Rupeshwar Sah [THA081BEI035]**
- **Laxman Shakya [THA081BEI016]**
- **Muhan Narshingh Shrestha [THA081BEI021]**

Department of Electronics and Computer Engineering
Institute of Engineering, Thapathali Campus
Tribhuvan University, Institute of Engineering, Thapathali Campus

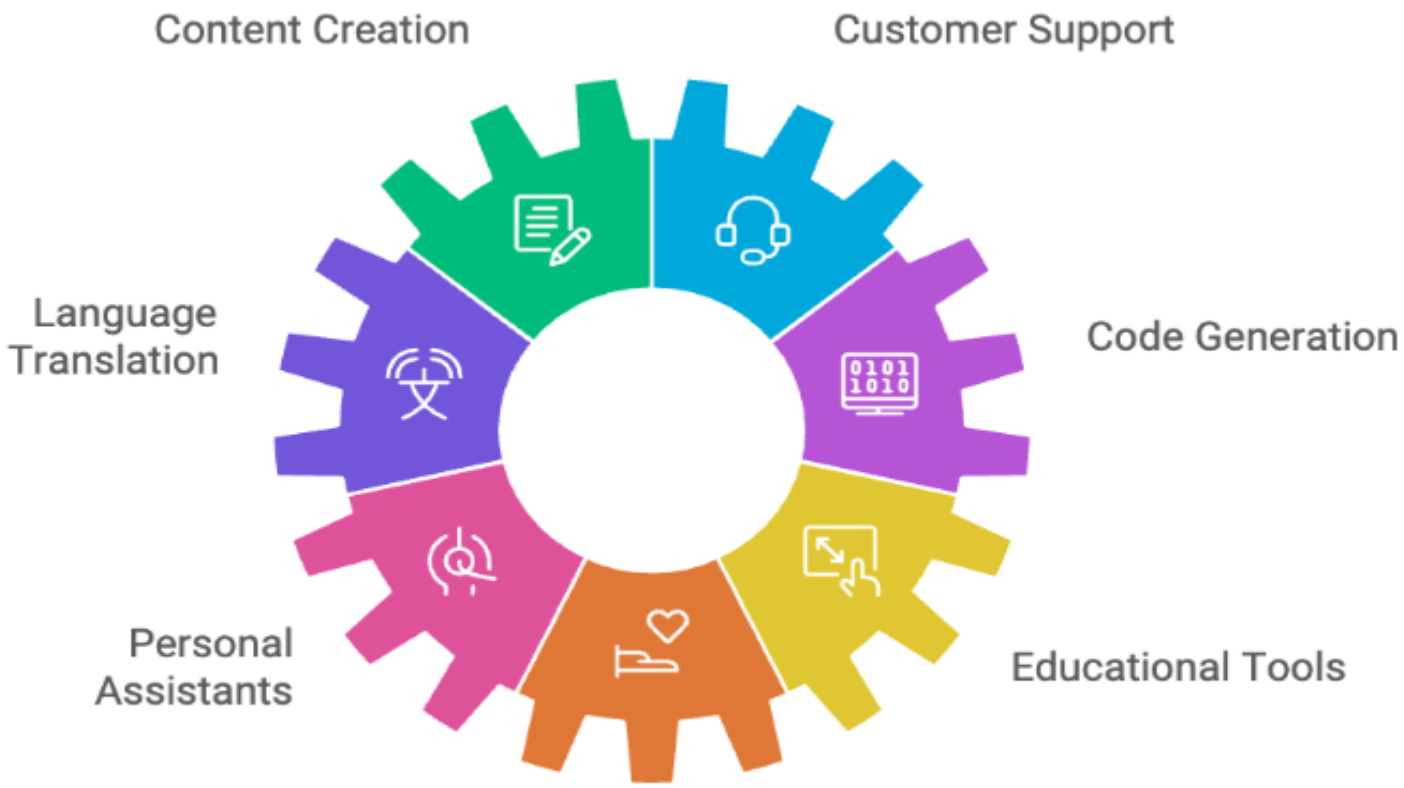
16 March 2025

Introduction to Large Language Models (LLMs)

What are LLMs?
Large Language Models (LLMs) are advanced AI systems trained on vast amounts of text data to understand, generate, and interact with human language.

Key Features:
- Understand and generate natural language
- Perform complex tasks like translation, summarization, and coding
- Continuously improve with more training and fine-tuning

Impact of LLMs on Different Industries



This Project Demonstrate

AI IN EDUCATION

WHY ?

Because our current education system lack…..

The Problem: Lack of Personalization

Traditional Methods

- Provide static content that doesn't adapt to individual learning styles.
- Depend on one-size-fits-all approaches like lectures and textbooks.
- Often lead to disengagement and inefficiency.

Digital Methods

- Include multimedia and interactivity but lack true personalization.
- Deliver pre-packaged content that rarely adjusts to unique student needs. needs.
- May cause information overload and reduced learning efficiency.

StudyBuddyy
The Personalization Solution

Study Notes & Quizzes

- Notes Generated based on user-specified topics, structured for clarity.
- Custom multiple-choice questions to reinforce knowledge.

Mind Map

- Diagrams to visualize topic relationships and boost retention.

YouTube Video Summaries

- Concise overviews of video content to save time.

Requirement Analysis

StudyBuddyy relies on specific software and tools for optimal performance.

Key Requirements:

- **Operating System**: Windows (7, 8, 10, or 11).
- **C Compiler**: GCC for compiling the C source code.
- **Python 3.x**: Runs content generation scripts.
- **Python Package**: learning-toolbox (v0.2.0) for AI-driven content creation.
- **Node.js & npm**: Powers mind map generation via Mermaid CLI.
- **Internet Connectivity**: Accesses external APIs like Gemini.

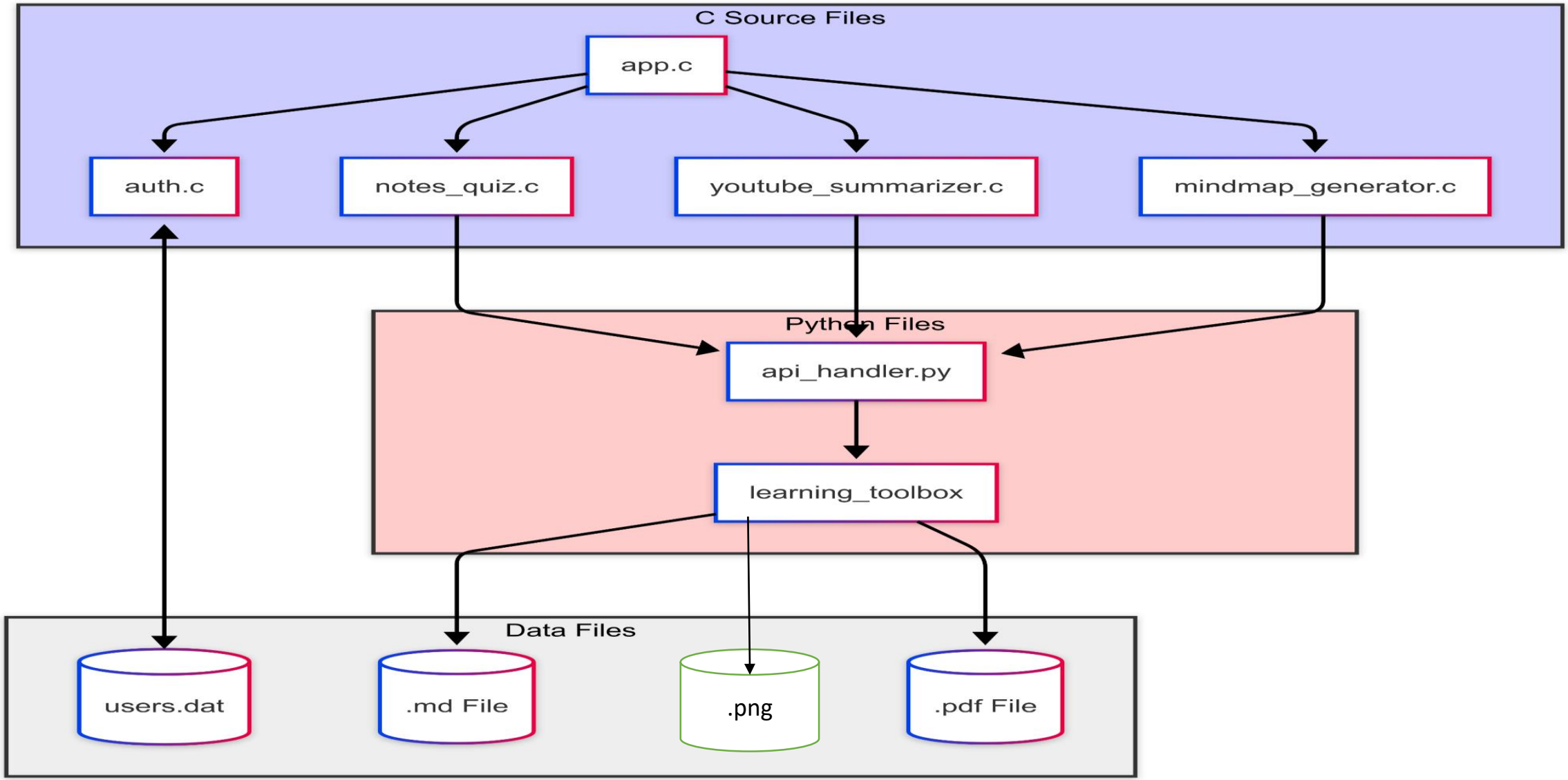
System Design and Architecture

StudyBuddyy uses a modular architecture for simplicity and maintainability.

Key Components:

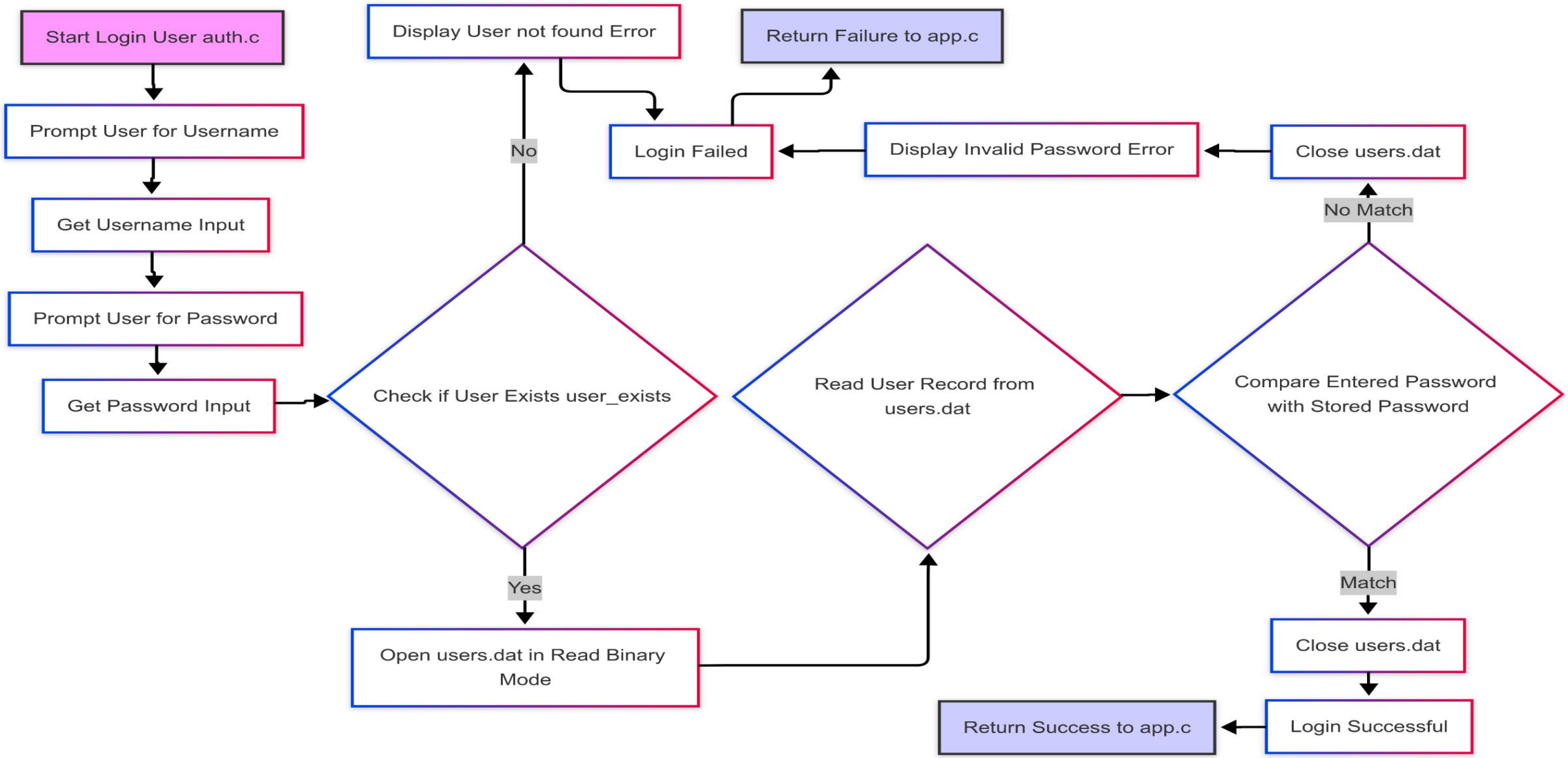
- **C Application (study_buddy.exe)**: Manages the user interface, authentication, and menu navigation.
- **Python API Handler (api_handler.py)**: Connects the C app to external AI APIs for content generation.
- **External APIs**: Gemini API via learning-toolbox for notes, quizzes, summaries, and mind maps.
- **Data Storage**: users.dat for authentication; file system for generated content.

Workflow: The C app processes user inputs, calls the Python handler to fetch AI-generated content, and saves results as files.



System Design and Architecture

auth.c Implementation



Implementation Details

Core Structure:

- **C Modules**:
  - **app.c:** Controls the main flow, UI, and feature invocation.
  - **auth.c:** Manages user login and registration.
  - **notes_quiz.c, youtube_summarizer.c, mindmap_generator.c:** Execute specific features via the Python handler.

- **Python Integration**:
  - **api_handler.py:** Processes requests from the C app, interacts with the AI API, and saves outputs.
  - Utilizes **learning-toolbox** for content generation.

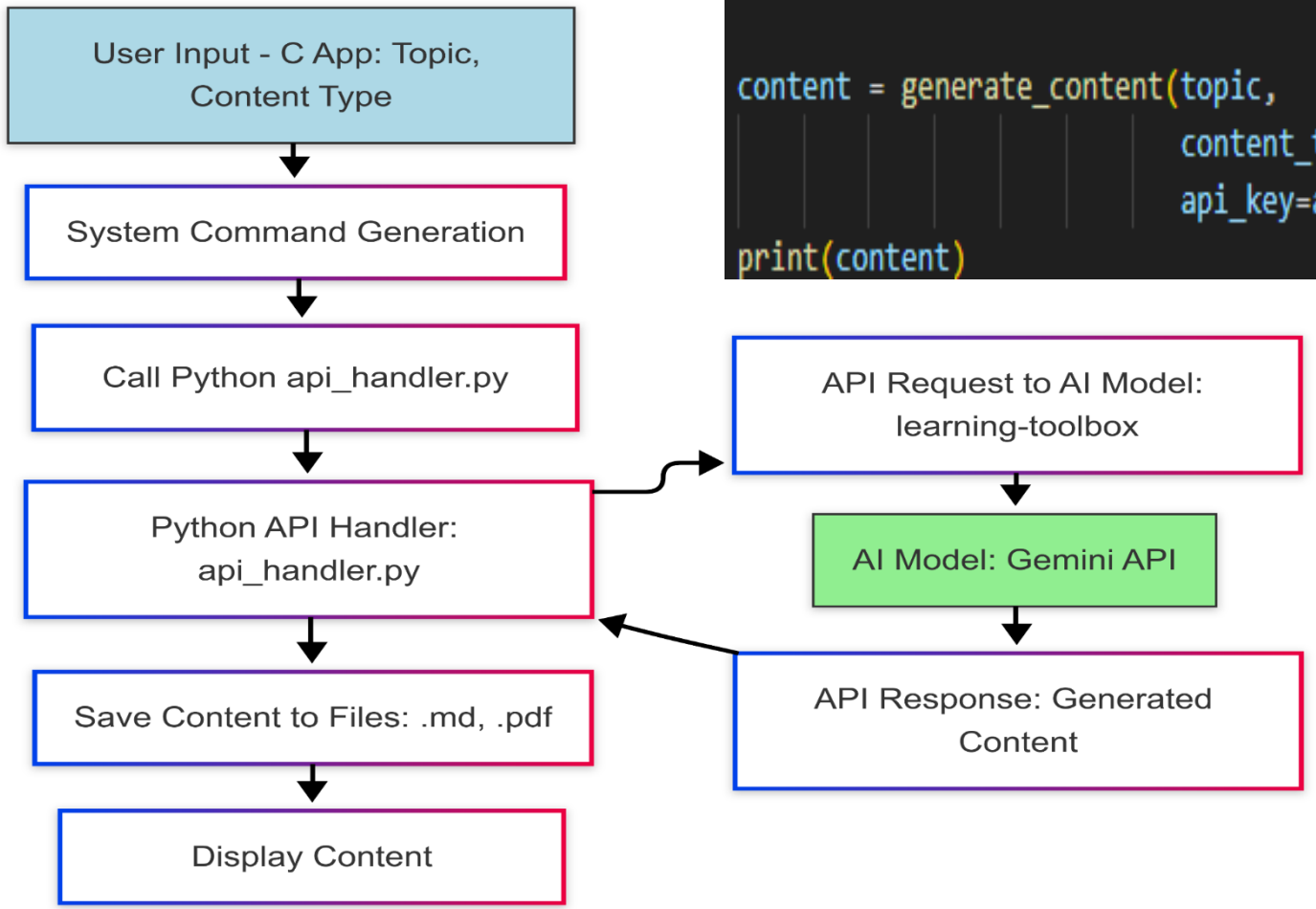
- **Key Techniques**:

1. System calls (system()) to run Python scripts from C.
2. File I/O for user data and content storage.

```c
// Call Python script
char cmd[512];
snprintf(cmd, sizeof(cmd),
"python api_handler.py \"%s\" %s",
topic, content_type);

system(cmd);
```

1




```python
topic = sys.argv[1]
content_type = sys.argv[2]

content = generate_content(topic,
                           content_type,
                           api_key=api_key)
print(content)
```

2

```python
save_files(topic,
           content,
           content_type,base_dir="")
```

3

Results and Analysis

Deeplearning

Overview

Deep Learning is a subfield of Machine Learning that is inspired by the structure and function of the human brain specifically, artificial neural networks. It's characterized by the use of **deep neural networks**, which are neural networks with multiple layers (typically more than three). These deep networks are capable of automatically learning intricate patterns and hierarchical representations from vast amounts of data without explicit programming of rules. This ability to learn complex features directly from raw data has led to breakthroughs in various fields like computer vision, natural language processing, speech recognition, and more.

In essence, Deep Learning automates feature engineering, a traditionally manual and time-consuming process in classical machine learning. Instead of hand-crafting features, deep learning algorithms learn optimal features directly from the data during the training process. This makes them incredibly powerful for tackling complex problems where feature extraction is not straightforward or intuitive.

Key Concepts

Here are some fundamental concepts crucial to understanding Deep Learning:

1. Neural Networks:

- **Structure:** At the heart of Deep Learning are neural networks. They are composed of interconnected nodes or neurons organized in layers. A typical neural network consists of:
  - **Input Layer:** Receives the raw input data. The number of neurons in this layer corresponds to the number of features in your input data.
  - **Hidden Layers:** One or more layers between the input and output layers. These layers perform the complex computations and feature extraction. "Deep" learning refers to networks with multiple hidden layers.
  - **Output Layer:** Produces the final output or prediction. The number of neurons in this layer depends on the task (e.g., 1 for regression, number of classes for classification).

Study Notes: Deep Learning notes with overview, concepts, and questions.

Analysis:

1. Structure and Formatting
2. Content Relevance and Accuracy
3. Usefulness for Learning

Potential Limitations:

1. **Depth of Information**
2. **Accuracy and Completeness**
3. **Over-reliance on AI**

Here are 10 MCQs about Machine Learning in markdown format:

1. **Which of the following is NOT a type of Machine Learning?**

   A. Supervised Learning B. Unsupervised Learning C. Reinforcement Learning D. Deep Learning

2. **What is the purpose of a validation set in Machine Learning?**

   A. To train the model B. To evaluate the final performance of the model after training C. To tune hyperparameters and prevent overfitting during training D. To deploy the model to production

3. **Which algorithm is best suited for classification problems?**

   A. Linear Regression B. K-Means Clustering C. Logistic Regression D. Principal Component Analysis

4. **What is 'overfitting' in Machine Learning?**

   A. When a model performs poorly on both training and test data B. When a model performs well on training data but poorly on test data C. When a model performs well on both training and test data D. When a model is too simple to capture the underlying pattern in the data

5. **Which of the following evaluation metrics is most appropriate for imbalanced datasets?**

   A. Accuracy B. Precision and Recall C. R-squared D. Mean Squared Error

6. **What is feature scaling?**

   A. Selecting the most important features from the dataset B. Transforming features to have a similar range of values C. Creating new features from existing features D. Removing irrelevant features from the dataset

7. **Which of the following is an example of unsupervised learning?**

   A. Spam email detection B. Image classification C. Customer segmentation D. Stock price prediction

Quizzes: Custom questions to test knowledge (e.g., "Machine Learning" quiz on algorithms).

Analysis:

1. Question Format
2. Relevance to Topic
3. Usefulness for Learning

Potential Limitations:

1. **Question Quality and Depth**
2. **Question Types**

Detailed Video Summary

Key Points

- **Colliding Blocks and Pi:** Two blocks colliding on a frictionless plane can be used to compute digits of pi based on the number of collisions.
- **Mass Ratio and Pi Digits:** As the mass ratio of the larger block to the smaller block increases by powers of 100, the number of collisions approximates pi to more digits.
- **Idealized Physics Puzzle:** This is presented as an idealized classical physics puzzle, ignoring real-world factors like energy loss, sound, and relativistic effects for large mass ratios.
- **Connection to Quantum Computing:** Surprisingly, this classical physics puzzle is secretly connected to quantum computing, specifically Grover's Algorithm for search.
- **Unsolved Problem Aspect:** The full connection to pi in this context is technically an unsolved mathematical problem, related to the digits of pi and potential off-by-one errors.
- **Problem-Solving Principles:** The video uses this puzzle to illustrate general problem-solving principles like listing relevant equations (conservation of energy and momentum), drawing pictures (state space), and respecting symmetries (transforming to a circular state space).
- **State Space and Geometry:** The problem is translated from physics into a geometric problem in a state space (velocity space), making it easier to visualize and solve.
- **Inscribed Angle Theorem:** The inscribed angle theorem from geometry is used to explain why the arcs in the state space diagram are of equal size, linking to the collision count.
- **Small Angle Approximation:** The small angle approximation ($\arctan(x) \approx x$ for small x) is crucial in connecting the mass ratio, the angle in the geometric representation, and the digits of pi.
- **Next Video: Quantum Computing Connection:** The video is the first part of a two-part series, with the next video explaining the connection to Grover's Algorithm and quantum computing.

Main Ideas

YouTube Summaries: Concise key points (e.g., summary of a physics video).

Analysis:

1. Conciseness

2. Key Point Extraction

3. Overall Summary

Potential Limitations:

1. **Loss of Nuance and Detail**
2. **Dependency on Transcript Quality**

Mind Maps: Visual topic relationships (e.g., a mindmap for a webdevelopment).

Analysis:

1. Visual Organization
2. Relationship Mapping
3. Usefulness for Learning

Potential Limitations:

1. **Complexity of Topics**
2. **Accuracy of Relationships**
3. **Customization and Interactivity**

Conclusion

Achievements:

- Integrates C and Python to build a functional AI study tool.
- Provides personalized materials that boost engagement and efficiency.

Significance:

- Showcases AI’s potential to revolutionize education through personalization.
- Saves time, enabling deeper understanding and critical thinking.

Future Work

Enhancements:

Security: Add password hashing and switch to SQLite for data storage.

UI: Develop a graphical interface for improved usability.

Features: Include flashcards, interactive exercises, and more content options.

Cross-Platform Support: Expand to macOS and Linux.

Personalization: Add user profiles to track learning styles and progress.

Vision

Transform StudyBuddyy into a versatile, comprehensive learning tool

AI could revolutionize education within the next five to ten years by delivering personalized content to a student's learning style and motivations.
- Bill Gates