**Flow diagram**

A diagram of a software system

Description automatically generated

**Component diagraA screenshot of a computer program

Description automatically generatedm**

**Scope**

**Fielder Recognition**: Recognising running or walking fielders in the outfield of the ground of a day/night cricket T20i match during the last 4 overs.

**Stat Genaration for the recognised fielder**:

Basic stat : Matches, Dismissals, Catches

Specific stat : user question islimited upto only 3 conditions/filters per question

**Features for the Face Recognition Model**

1. **Face Detection:**

To detect faces within images or video frames

1. **Feature Extraction (Embeddings):**

To convert each detected face into a compact, high-dimensional feature vector

1. **Training the Classifier:**

To classify embeddings into identities based on labelled training data

**Features for the Hybrid Spatio-Temporal Gait analysis Model**

**Spatial Features:**

1. **Jersey Numbers and Player Names**:

As each player has a unique number and name on their jersey.

1. **Jersey Colours and Country Names**:

help to distinguish players from different teams.

1. **General fielding position**:

Long-on, Long-off, Deep cover, Deep extra cover, Deep point, Deep backward point, Deep third man, Deep fine leg, Deep backward square leg, Deep square leg, Deep mid-wicket

**(**total of **11** positions)

**Temporal Features:**

1. **Gait Cycle Features**: **Step frequency, Step length, Symmetry**.

Rhythm and style of the player movement

1. **Kinematic Features**: **Joint angles, Limb velocities.**

unique biomechanical signatures

**Research Questions**

**Research Question 1: Player Recognition**

*“How effectively can a computer vision ensemble model employing face recognition and gait analysis using spatio-temporal features, recognise fielders in the outfield during the last four overs of a day/night T20 International cricket match?”*

**Aim:**

To develop and validate a robust computer vision ensemble model that effectively employs face recognition and spatio-temporal gait analysis for accurate fielder recognition during the critical final overs of T20 International cricket matches, enhancing the analytical capabilities under varying environmental conditions.

**Objectives:**

1. **Develop an Ensemble Model:** Construct a model that leverages both facial recognition and hybrid spatio-temporal gait analysis to identify fielders accurately. This model will integrate various analytical techniques to capture static and dynamic attributes of players.
2. **Data Collection and Preprocessing:** Gather a comprehensive dataset encompassing video feeds and images that reflect diverse lighting conditions and field settings typical of T20 matches. This data will be preprocessed to normalize variations and prepare for effective model training.
3. **Feature Engineering:** Identify critical features that significantly impact player recognition. This includes detailed analysis of gait patterns, player jerseys, and contextual environment settings.
4. **Model Training and Optimization:** Train the ensemble model using advanced machine learning techniques. Optimize the model for high accuracy and operational efficiency in dynamic, real-time match conditions.
5. **Validation under Various Conditions:** Evaluate model performance across different scenarios to ensure reliability and robustness. This involves testing under variable lighting, different player movements, and multiple field positions.

**Research Question 2: Stat Generation for Recognized Players**

*“How can a large language model be prompt-engineered to accurately translate natural language user-defined questions, with up to three conditions, into SQL queries for generating accurate and relevant statistics?”*

**Aim:**

To harness the capabilities of a prompt-engineered large language model for translating complex natural language queries into precise SQL queries, thereby enabling the generation of accurate and contextually relevant statistics for recognized players in T20 cricket matches, with an emphasis on optimizing interaction and query responsiveness under varying analytical demands.

**Objectives:**

1. **Selection of an Optimal Large Language Model (LLM)**:

Choose the most suitable LLM from available options such as ChatGPT, Gemini, and Lama, based on their capabilities to understand and process natural language queries effectively.

2. **Design of a Relational Database**:

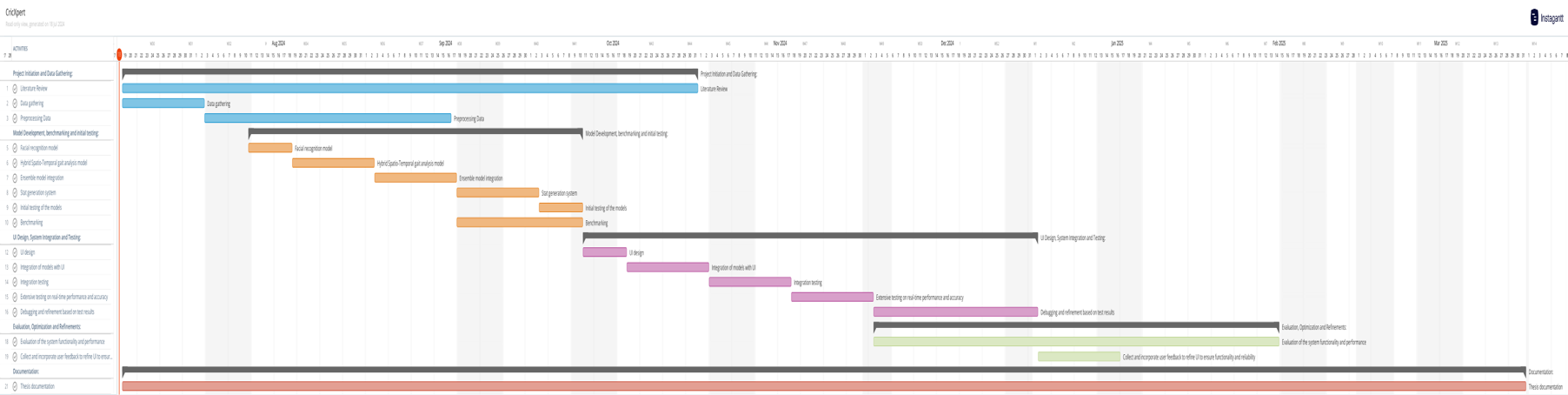
Construct a robust and scalable relational database designed to store detailed player statistics. This database should support efficient real-time SQL queries and be capable of handling user-defined filters and recognition inputs, facilitating the retrieval of specific player data as required.

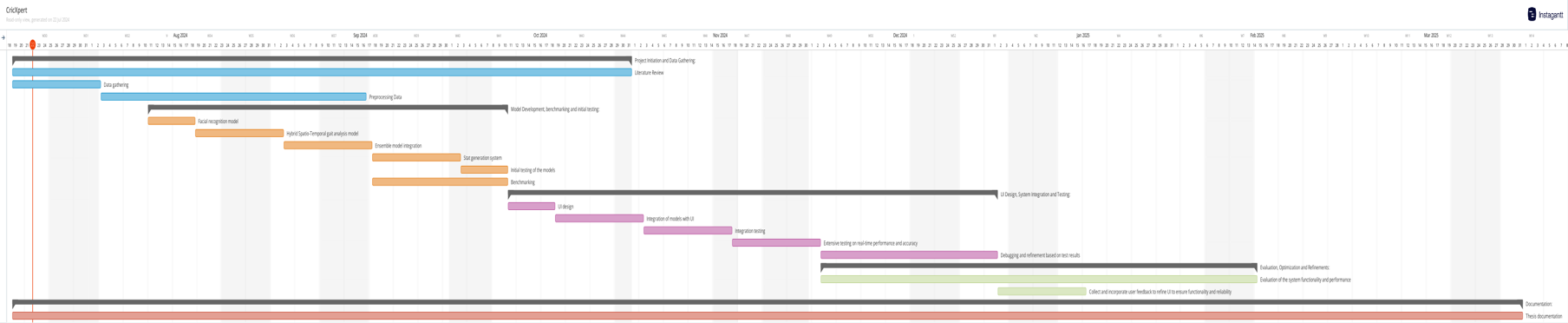
3. **Implementation of Prompt Engineering Techniques**:

Apply and refine prompt engineering strategies to train the selected LLM on accurately interpreting natural language queries. Optimize these prompts to improve the LLM’s ability to formulate SQL queries that are both syntactically correct and logically consistent with user intentions.

4. **Evaluation of Input Character Length on Filter Effectiveness**:

Assess how different filter configurations and the lengths of input characters affect the relevance and accuracy of the generated statistics, aiming to optimize the system for various analytical requirements.

**Timeline of the project**



**Evaluation Metrics for Player Recognition (Gait Model)**

**Spatial Component (**CNN**)**

**F1 Score**: provides a critical balance between precision and recall, essential for accurate player analytics where both identifying every player correctly and minimizing incorrect identifications are equally important

1. **Powers, D.M.W. (2011)**. “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation.”This journal article elaborates on the importance of F1 Score in providing a more truthful measure of a model’s predictive power, particularly useful in your scenario for balancing the identification accuracy of players.
2. **Sokolova, M., & Lapalme, G. (2009)**. “A systematic analysis of performance measures for classification tasks.” This paper explores various metrics and validates the use of F1 Score for its ability to harmonize precision and recall, making it ideal for tasks where both false positives and false negatives carry significant weight.

**Mean Average Precision (mAP)**: offers a robust method for evaluating performance across multiple players and recognition thresholds, which aligns well with the multi-player tracking.

1. **Everingham, M., et al. (2010)**. “The Pascal Visual Object Classes (VOC) Challenge. “This paper discusses the use of mAP for evaluating the performance of object detection algorithms, underscoring its effectiveness in handling multiple classes and varying thresholds of detection, which is analogous to recognizing multiple players in a sports scenario.
2. **Lin, T.-Y., et al. (2014)**. “Microsoft COCO: Common Objects in Context.”

mAP is highlighted as a key metric for comprehensive evaluation across different object detection models, supporting its use in scenarios with complex backgrounds and multiple object classes—similar to differentiating players from various teams.

**Temporal Component (**RNN, LSTM, or GRU**)**

**Sequence Accuracy**: It assesses whether the entire sequence is recognized as belonging to one player, crucial for continuous tracking in sports analytics.

1. **Graves, A., et al. (2013)**. “Hybrid Speech Recognition with Deep Bidirectional LSTM.” In this paper, the importance of sequence accuracy is highlighted in the context of speech recognition using LSTM networks, where maintaining the sequence integrity is crucial. This concept can be adapted to player recognition in videos, emphasizing the need for accurate sequence output over time.
2. **Zhang, K., Zhang, L., & Yang, M-H. (2012)**. “Real-Time Object Tracking via Online Discriminative Feature Selection.” This paper, although focused on object tracking, discusses the importance of accurate sequential tracking in video, which aligns with the need for sequence accuracy in tracking player movements based on gait and other features.

**Edit Distance (Levenshtein Distance)**: how closely a predicted sequence of movements matches the actual movement sequence observed in video footage

1. **Sakoe, H., & Chiba, S. (1978)**. “Dynamic Programming Algorithm Optimization for Spoken Word Recognition.” This seminal paper introduces the use of dynamic programming to compute distances between sequences, which forms the basis for using Edit Distance in temporal sequence evaluation. It is highly applicable for validating movement sequences in sports analytics.
2. **Levenshtein, V.I. (1966)**. “Binary codes capable of correcting deletions, insertions, and reversals.” This foundational paper describes the metric itself, and its applications have extended into numerous fields, including bioinformatics for sequence alignment, which parallels movement sequence analysis in sports technology.

**Evaluation Baselines for Player Recognition**

1. **Temporal Model Only**: Implement an RNN, LSTM, or GRU model to process only the temporal aspects of the data, serving as a baseline to evaluate the effectiveness of sequence modelling without spatial features.

2. **Spatial Model Only**: Deploy a CNN to analyse solely the spatial data from static images, providing a baseline to measure the identification capabilities of spatial features alone.

3. **Conventional Machine Learning Models**: Use traditional machine learning models like SVM and Random Forests, which require engineered features, to compare the efficacy of simpler methodologies against more complex deep learning approaches.

4. **Other Deep Learning Architectures**: Compare with advanced Transformer models specialized in video and motion data analysis, to assess the performance of the proposed model.

**Evaluation Metrics for Stat Generation**

**Accuracy of the SQL Queries:** ensures the model meets the basic requirement of correctly interpreting and translating user input into SQL

1. **Zhong, V., Xiong, C., & Socher, R. (2017).** “Seq2SQL: Generating Structured Queries from Natural Language using Reinforcement Learning.” This paper introduces a model that translates natural language queries into SQL, emphasizing the importance of accuracy in query generation for the effectiveness and reliability of such systems.
2. **Yu, T., et al. (2018).** “Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task.” The Spider benchmark emphasizes accuracy as a critical metric for evaluating the capability of models to generate correct SQL queries from complex and diverse natural language descriptions across various domains.

**Query Handling Complexity**: measures a model’s ability to handle complex SQL queries, including those with multiple conditions, joins, and subqueries

1. **Wang, C., et al. (2020).** “RAT-SQL: Relation-Aware Schema Encoding and Linking for Text-to-SQL Parsers.” This paper discusses the challenges and solutions in handling complex SQL query generation from natural language, specifically addressing the need for models to manage intricate relational structures and operators.
2. **Lin, X., Socher, R., & Xiong, C. (2020).** “Bridging Textual and Tabular Data for Cross-Domain Text-to-SQL Semantic Parsing.” This research further validates the need for handling complexity in SQL query generation, providing insights into how models can better understand and integrate context from both text and structured data to form accurate and complex SQL queries.

**Response Time**: Time taken from the final input of the user query to the retrieval of results.Critical in real-time applications where speed of data retrieval is as important as accuracy.

**Evaluation Baselines for Stat Generation**

1. **Basic Prompt-Engineered LLM**: Implement a baseline large language model with basic prompt engineering to evaluate its fundamental ability to translate natural language questions into SQL queries.

2. **Comparison with Other LLMs**: Assess with other advanced large language models that have been optimized for SQL generation tasks.

**Best LLM for Stat Generation**

**ChatGPT 4**

1. **(‘CHESS: Contextual Harnessing for Efficient SQL Synthesis’ 2024)**. This study introduces a system that efficiently synthesizes SQL queries from natural language using a scalable LLM-based pipeline. It highlights the importance of integrating contextual information from database schemas and values, demonstrating that LLMs can effectively handle complex real-world database scenarios by retrieving and utilizing relevant information efficiently .
2. **(Guo et al. 2024)** **Retrieval-Augmented GPT-3.5-Based Text-to-SQL Framework** **with Sample-Aware Prompting and Dynamic Revision Chain**

This research discusses a framework that enhances text-to-SQL conversion by incorporating intermediate representations and retrieval-augmented strategies. It underscores the capability of GPT models to adapt to cross-domain databases, thus validating the potential of similar technologies like GPT-4 for your project.

1. **(‘Evaluating the Text-to-SQL Capabilities of Large Language Models’ 2024).** This paper focuses on analyzing the performance of LLMs, specifically Codex (based on GPT-3), in generating SQL queries. It provides insights into the effectiveness of using LLMs for text-to-SQL tasks, discussing both the strengths in generating contextually accurate SQL queries and the limitations, such as handling complex query structures .

Spatial Model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Regularization | Data Augmentation | F1 Score | mAP | Remarks |
| A basic CNN with 7 layers | - | NO | 98 | 98 | overfitted |
| A basic CNN with 7 layers |  |  |  |  |  |
| MobileNetV2 | - | NO | 92 | 97 | overfitted |
| MobileNetV2 | - | YES | 68 | 81 | Not performing |
| MobileNetV2 | Dropout | YES | 51 | 59 | Not performing |
| EfficientNetB0 | Dropout | NO | 100 | 100 | overfitted |
| EfficientNetB0 | Dropout | YES | 99 | 99 | shape |
| EfficientNetB0 | Dropout, L2 | YES | 98 | 98 | overfitted |
| ResNet50 | Dropout | YES | 75 | 89 | Not performing |
| ResNet50 | Dropout | NO | 98 | 98 | shape |
| DenseNet121 | Dropout | NO | 75 | 86 | Not performing |
| DenseNet121 | Dropout | YES | 51 | 60 | Not performing |
| InceptionV3 | Dropout | NO | 51 | 61 | Not performing |
| ViT |  |  |  |  |  |

1. Automatic Player Face Detection and Recognition for Players in Cricket Games - IEEE Access

2. Automatic Player Detection and Identification for Sports Entertainment Applications - Pattern Analysis and Applications

3. Enhancing Cricket Performance Analysis with Human Pose Estimation and Machine Learning - Italian National Conference on Sensors

4. A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions - Applied Sciences

5. Optimized Deep Learning-Based Cricket Activity Focused Network - Alexandria Engineering Journal

6. Cricket Shot Detection Using 2D CNN - ICICCS

7. Multi-camera Multi-player Tracking with Deep Player Identification - Pattern Recognition

8. Optical Tracking in Team Sports - Journal of Quantitative Analysis in Sports

9. Automated Recognition of the Cricket Batting Backlift - Scientific Reports

10. Cricket Scoreboard Automation using Umpire Gestures - IJRESM

11. Past, Present, and Future of Face Recognition: A Review - Electronics

12. Hybrid Deep-Learning Framework Based on Gaussian Fusion of Multiple Spatiotemporal Networks for Walking Gait Phase Recognition

13. HybridGait: A Benchmark for Spatial-Temporal Cloth-Changing Gait Recognition with Hybrid Explorations

14. Gait-DenseNet: A Hybrid Convolutional Neural Network for Gait Recognition.

15. An Acceleration Based Fusion of Multiple Spatiotemporal Networks for Gait Phase Detection

16. Robust gait recognition using hybrid descriptors based on Skeleton Gait Energy Image

17. Model-based person identification in multi-gait scenario using hybrid classifier

18. Hybrid Deep Neural Network Framework Combining Skeleton and Gait Features for Pathological Gait Recognition

19. Gaitcotr: Improved Spatial-Temporal Representation for Gait Recognition with a Hybrid Convolution-Transformer Framework

20. Multi-view gait recognition system using spatio-temporal features and deep learning

21. Gait Recognition Analysis for Human Identification Analysis-A Hybrid Deep Learning Process

22. STAR: Spatio-Temporal Augmented Relation Network for Gait Recognition

23. Automatic multi-gait recognition using pedestrian’s spatiotemporal features

24. Human Gait Recognition: A Single Stream Optimal Deep Learning Features Fusion

25. Skeleton-based abnormal gait recognition with spatio-temporal attention enhanced gait-structural graph convolutional networks

26. Hybrid LSTM and GAN model for action recognition and prediction of lawn tennis sport activities

27. INDIVIDUAL ACTION AND GROUP ACTIVITY RECOGNITION IN SOCCER VIDEOS

28. Identification of humans using gait

29. Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: a preliminary analysis with multiple trackers, viewing angles, and walking directions

30. Accuracy of Computer Vision-Based Pose Estimation Algorithms in Predicting Joint Kinematics During Gait

31. Two-dimensional video-based analysis of human gait using pose estimation

32. Gait recognition using spatio-temporal silhouette-based features

33. Ubiquitous Gait Analysis through Footstep-Induced Floor Vibrations

34. Quantitative and Qualitative Running Gait Analysis through an Innovative Video-Based Approach

35. Gait symmetry methods: Comparison of waveform-based Methods and recommendation for use

36. Sensor-Based Human Activity Recognition with Spatio-Temporal Deep Learning

37. Multimodal Low Resolution Face and Frontal Gait Recognition from Surveillance Video

38. Real-Time Human Recognition at Night via Integrated Face and Gait Recognition Technologies

39. Multimodal Adaptive Fusion of Face and Gait Features using Keyless attention based Deep Neural Networks for Human Identification

40. Gait Recognition by Jointing Transformer and CNN

41. Exploring Deep Models for Practical Gait Recognition

42. GaitPT: Skeletons Are All You Need For Gait Recognition

43. A model-based gait recognition method with body pose and human prior knowledge

44. Human gait recognition: A systematic review

45. Gait analysis for recognition and classification

46. Human Recognition by Appearance and Gait

47. Enhancing Natural Language Query to SQL Query Generation Through Classification-Based Table Selection

48. Structure Guided Large Language Model for SQL Generation

49. MCS-SQL: Leveraging Multiple Prompts and Multiple-Choice Selection For Text-to-SQL Generation

50. Knowledge-to-SQL: Enhancing SQL Generation with Data Expert LLM

51. Prompting GPT-3.5 for Text-to-SQL with De-semanticization and Skeleton Retrieval

52. RH-SQL: Refined Schema and Hardness Prompt for Text-to-SQL

53. Enhancing Text-to-SQL Capabilities of Large Language Models: A Study on Prompt Design Strategies

54. Next-Generation Database Interfaces: A Survey of LLM-based Text-to-SQL

55. Bridging Language & Data: Optimizing Text-to-SQL Generation in Large Language Models

56. CodeS: Towards Building Open-source Language Models for Text-to-SQL

57. [2405.16755] CHESS: Contextual Harnessing for Efficient SQL Synthesis (arxiv.org)

58. Retrieval-Augmented GPT-3.5-Based Text-to-SQL Framework with Sample-Aware Prompting and Dynamic Revision Chain

59. Evaluating the Text-to-SQL Capabilities of Large Language Models

*“How accurately can an ensemble model that integrates facial recognition and hybrid spatio-temporal gait analysis differentiate between fielders in a T20 cricket match setting, considering variables such as lighting conditions, player movement dynamics, and field positions?”*

**“How effectively can a computer vision model, tailored for dynamic environments, identify and track fielders based on their gait and facial features during the critical last four overs of a day/night T20 International cricket match, considering variations in lighting, player speed, and field positioning?”**

1. This research is primarily about recognising of humans/players

2. Recognising running or walking fielders in the outfield of the ground of a day/night cricket T20i match during the last 4 overs.

3. This uses an ml model that combines face recognition and Gait analysis(Spatio temporal)

4. Usage of Spatio temporal features for human recognition instead of human activity recognition (like kind of transfer learning)

*“How effectively can a system generate stats for individual cricket players recognized through video analysis, and how do user-defined conditions (up to three filters) influence the accuracy and relevance of these statistics during a T20 match?”*

*How to effectively can a system generate relevant and accurate statistics by user-defined conditions (up to three filters) for a T20 cricket player?*

*How to generate relevant and accurate statistics for natural written user-defined question with conditions (up to 3) by using prompt engineered LLM to convert text to SQL query?*