VIVEKANAND EDUCATION SOCIETY'S
INSTITUTE OF TECHNOLOGY

Department of Computer Engineering

Project Report on

GameSpec Advisor

In partial fulfillment of the Fourth Year (Semester–VII), Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai

Academic Year 2024-2025

Prof. Mr. Richard Joseph

Submitted by

Ayush Rakesh Gerra / 18
Anisha Sunil Rohra / 52
Manan Dadlani / 09
Meet Kewalramani / 29

(2024-25)

VIVEKANAND EDUCATION SOCIETY'S
INSTITUTE OF TECHNOLOGY

Department of Computer Engineering

CERTIFICATE of Approval

This is to certify that **Ayush Rakesh Gerra (18), Anisha Sunil Rohra (52), Manan Dadlani (09), Meet Kewalramani (29)** of Fourth Year Computer Engineering studying under the University of Mumbai has satisfactorily presented the project on "**GameSpec: Advisor**" as a part of the coursework of PROJECT-I for Semester-VII under the guidance of professor **Richard Joseph.** in the year 2024-2025.

__23/10/2024__
Date

Internal Examiner

External Examiner

Project Mentor
Richard Joseph

Head of the Department
Dr. Mrs. Nupur Giri

Principal
Dr. J. M. Nair

ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to Assistant Professor **Mr. Richard Joseph** (Project Guide) for his kind help and valuable advice during the development of project synopsis and for his guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr. (Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair ,** for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is a great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

Computer Engineering Department

COURSE OUTCOMES FOR B.E PROJECT

Learners will be to:-

Course Outcome	Description of the Course Outcome
CO 1	Do literature survey/industrial visit and identify the problem of the selected project topic.
CO2	Apply basic engineering fundamental in the domain of practical applications FORproblem identification, formulation and solution
CO 3	Attempt & Design a problem solution in a right approach to complex problems
CO 4	Cultivate the habit of working in a team
CO 5	Correlate the theoretical and experimental/simulations results and draw the proper inferences
CO 6	Demonstrate the knowledge, skills and attitudes of a professional engineer & Prepare report as per the standard guidelines.

Abstract

Building a budget gaming PC that optimally balances performance and cost can be a challenging task, especially with constantly changing hardware options and pricing. The proposed system aims to simplify this process by offering a personalized platform that provides real-time hardware recommendations, compatibility checks, and performance analysis tailored to the user's budget, gaming preferences, and performance goals.

The system utilizes web scraping to gather data from gaming platforms like Steam and hardware sites such as PCPartPicker, and integrates sentiment analysis from user reviews to identify common issues or strengths in hardware configurations. By implementing advanced techniques such as feature correlation, association rule mining, and benchmark-based performance analysis, the system dynamically recommends the best hardware setups for optimal gaming performance.

In addition to providing real-time hardware price comparisons, the system offers a comprehensive PC health check that analyzes the user's current setup, identifies bottlenecks, and suggests upgrades or optimizations. Driver recommendations are also provided to ensure optimal system stability and performance. The user-friendly interface allows users to easily input their preferences and receive tailored recommendations, ensuring that both novice and experienced gamers can build high-performance gaming PCs within their budget constraints.

By leveraging advanced data processing, real-time benchmarking, and personalized insights, this system provides a holistic solution for budget-conscious gamers looking to optimize their gaming experience.

INDEX

Chapter 1 : Introduction

1.1. Introduction

The gaming industry is experiencing unprecedented growth, driving the demand for highly optimized, custom gaming setups. As gamers and tech enthusiasts seek to enhance their experience with the best possible hardware configurations, they face challenges in navigating vast and often fragmented information about system requirements, hardware compatibility, and emerging trends. This project aims to address these challenges by developing a comprehensive platform that not only consolidates data from multiple sources but also applies advanced analytical techniques to extract actionable insights. The platform will enable users, from beginners to seasoned gamers, to make informed decisions, optimizing their systems for the latest games while future-proofing their setups. By streamlining the complex process of matching system requirements with hardware capabilities, this tool will become an indispensable resource in the evolving gaming landscape.

1.2. Motivation

The motivation for this project is driven by several key factors that highlight the need for a comprehensive platform for gaming setups:

1. **Empowering Gamers**: The primary motivation is to empower gamers, particularly those who may lack technical expertise, by providing them with the knowledge and tools necessary to make informed decisions about their gaming hardware. By simplifying the process of identifying and configuring optimal setups, the project aims to enhance the overall gaming experience and encourage more individuals to engage with this dynamic community.
2. **Addressing Complexity and Confusion**: With the vast array of gaming components and configurations available, many users feel overwhelmed and confused. The motivation lies in alleviating this complexity by creating a user-friendly platform that consolidates

critical information, offers compatibility analysis, and provides tailored recommendations.

3. **Staying Current with Industry Trends**: The gaming industry is rapidly evolving, with new hardware and technologies emerging frequently. The motivation for this project includes the desire to keep users informed about the latest trends and advancements, ensuring they can make future-proof decisions. By providing timely updates and insights, the platform aims to foster a sense of confidence and engagement among users.
4. **Encouraging Inclusivity in Gaming**: Another significant motivation is to create a more inclusive gaming environment by supporting newcomers as they enter the gaming world. The platform aims to demystify technical jargon, provide educational resources, and guide users through the complexities of hardware selection. This can help bridge the gap between seasoned gamers and novices, fostering a welcoming community for all.
5. **Enhancing Gaming Performance**: The ultimate goal is to enable users to optimize their gaming performance by configuring setups that are tailored to their specific needs and preferences. By analyzing system requirements and compatibility, the project seeks to enhance user satisfaction and enjoyment, allowing gamers to fully immerse themselves in their gaming experiences without the frustration of technical difficulties.

By addressing these motivations, the project aspires to create a valuable resource that not only supports the gaming community but also contributes to its growth and evolution.

1.3. Problem Definition

Configuring the ideal gaming setup has become increasingly complex, given the vast array of available components, configurations, and the rapid pace of technological advancements. Gamers, especially those new to the field, often find it overwhelming to sift through numerous hardware options, understand system requirements for different games, and navigate compatibility issues across processors, GPUs, memory, and other critical components. This complexity is further compounded by the constant evolution of gaming hardware trends, making it difficult to stay updated and select future-proof configurations.

There is a pressing need for a platform that offers a comprehensive, user-friendly solution. Such a platform should not only consolidate and analyze data from multiple sources but also provide personalized recommendations and in-depth compatibility assessments. This would empower users to make informed decisions, reducing the guesswork involved in building or upgrading gaming systems and optimizing performance for the latest games.

The platform must also serve as an educational resource, introducing neophytes to the exhilarating world of gaming. By demystifying complex hardware concepts and offering tailored guidance, it can ease the entry barrier for new gamers, ensuring they are equipped with the knowledge and tools to configure a setup that delivers an optimal gaming experience.

Key Problems:

1. **Difficulty in Identifying and Configuring the Right Gaming Setup**: The sheer number of available components and configurations creates confusion, making it hard for users to choose the optimal combination for their gaming needs.
2. **Lack of Comprehensive Analysis and Guidance**: Users need a platform that simplifies decision-making by providing thorough analysis and actionable insights, ensuring that both novices and experienced gamers can build systems with confidence.
3. **Introducing Neophytes to Gaming**: There is a need for an accessible platform that educates and guides newcomers, making their entry into gaming smoother while also providing insights into hardware, compatibility, and trends.

1.4. Relevance of the Project

The relevance of this topic stems from the explosive growth of the gaming industry, which has fueled demand for optimized, high-performance gaming setups. As gaming becomes more mainstream and hardware evolves at an unprecedented pace, the need for a comprehensive platform that can simplify decision-making is more critical than ever.

Currently, gamers face numerous challenges, including understanding the technical specifications of hardware, staying updated with the latest trends, and ensuring that their components are compatible for optimal performance. This is especially relevant as games continue to become more demanding, requiring specialized hardware configurations to run smoothly. For newcomers, these challenges are even more pronounced, as the complexity of hardware components and configurations can act as a barrier to entry into gaming.

The project is highly relevant as it addresses these pain points by providing a centralized, intelligent platform. This platform can not only educate and assist users in selecting the best hardware for their needs but also adapt to the rapidly changing technology landscape. It aims to consolidate fragmented information from various sources, provide personalized guidance, and perform in-depth compatibility analyses, making it a valuable tool for both casual gamers and tech enthusiasts.

In the context of increasing interest in gaming, the project's relevance is also tied to the growing community of individuals who seek accessible, reliable, and up-to-date information to optimize their gaming experience. By creating a user-friendly platform that simplifies hardware selection and configuration, this project has the potential to become a vital resource for a wide audience, from beginners exploring the world of gaming to experienced gamers looking to enhance their setups.

1.5. Methodology used

The methodology for this project begins with data extraction, where system requirements and game metadata are scraped from platforms like Steam, Epic Games, and Games of Galaxy using web scraping tools or APIs. After collecting the relevant hardware specifications and performance metrics, the next step is data processing. This involves extracting captions or text from the scraped data, performing sentiment analysis on user reviews, and identifying key features such as GPU, CPU, RAM, and storage requirements.

Following this, the configuration analysis phase looks at popular pre-built PC configurations and trending setups to understand which hardware combinations perform well. A feature correlation matrix is then built, mapping the relationships between different hardware components (e.g., CPU-GPU combinations) and how they impact gaming performance. This matrix helps in assessing how well different components work together.

In the compatibility analysis stage, the project utilizes data from GPU benchmarks, YouTube reviews, and hardware reviews, cross-referenced with the feature correlation matrix, to determine compatibility between various hardware setups and specific game requirements. Association rule mining is employed to discover common patterns and associations between hardware components, identifying frequently successful combinations in gaming setups.

Next, a PC health check evaluates the user's current system setup, identifying potential bottlenecks or areas for optimization, and providing recommendations for hardware upgrades or system tweaks. Driver recommendations follow, offering users the latest driver updates based on their system's hardware and operating system.

Finally, the website development stage involves creating a static website that provides users with personalized gaming setup guidance. This site allows users to input their system specifications and receive recommendations on hardware compatibility, driver updates, and potential optimizations for better gaming performance.

Chapter 2: Literature Survey

2.1. Research Papers

1. Comparative Analysis of Machine Learning Models for Performance Prediction of the SPEC Benchmarks
    a. **Abstract of the research paper:**

    Simulation-based performance prediction is cumbersome and time-consuming. An alternative approach is to consider supervised learning as a means of predicting the performance scores of Standard Performance Evaluation Corporation (SPEC) benchmarks. SPEC CPU2017 contains a public dataset of results obtained by executing 43 standardised performance benchmarks organised into 4 suites on various system configurations. This paper analyses the dataset and aims to answer the following questions: I) can we accurately predict the SPEC results based on the configurations provided in the dataset, without having to actually run the benchmarks? II) what are the most important hardware and software features? III) what are the best predictive models and hyperparameters, in terms of prediction error and time? and IV) can we predict the performance of future systems using the past data? We present how to prepare data, select features, tune hyperparameters and evaluate regression models based on Multi-Task Elastic-Net, Decision Tree, Random Forest, and Multi-Layer Perceptron neural networks estimators. Feature selection is performed in three steps: removing zero variance features, removing highly correlated features, and Recursive Feature Elimination based on different feature importance metrics: elastic-net coefficients, tree-based importance measures and Permutation Importance. We select the best models using grid search on the hyperparameter space, and finally, compare and evaluate the performance of the models. We show that tree-based models with the original 29 features provide accurate predictions with an average error of less than 4%. The average error of faster Decision Tree and Random Forest models with 10 features is still below 6% and 5% respectively.

b. Inference drawn from the paper:

1. SPEC CPU2017 Dataset: The authors utilize the SPEC CPU2017 dataset, which contains results from 43 standardized performance benchmarks organized into four suites.
2. Multi-Target Regression Models: The research explores multi-target regression techniques, which allow for the simultaneous prediction of multiple performance metrics.
3. Model Evaluation: The authors conduct a thorough evaluation of different ML models, including Multi-Task Elastic-Net, Decision Trees, Random Forests, and Multi-Layer Perceptrons.

2. Evaluating Machine Learning Models for Disparate Computer Systems Performance Prediction

a. Abstract of the research paper:

Performance prediction is an active area of research due to its applicability in the advancements of hardware-software co-development. Several empirical machine-learning models, such as linear models, non-linear models, probabilistic models, tree-based models and, neural networks, are used for performance prediction. Furthermore, the prediction model's accuracy may vary depending on performance data gathered for different software types (compute-bound, memory-bound) and different hardware (simulation-based or physical systems). We have examined fourteen machine-learning models on simulation-based hardware and physical systems by executing several benchmark programs with different computation and data access patterns. Our results show that the tree-based machine-learning models outperform all other models with median absolute percentage error (MedAPE) of less than 5% followed by bagging and boosting models that help to improve weak learners. We have also observed that prediction accuracy is higher on simulation-based hardware due to its deterministic nature as compared to physical systems. Moreover, in physical systems, the prediction accuracy of memory-bound algorithms is higher as compared to compute-bound algorithms due to manufacturer variability in processors.

b. Inference drawn from the paper:

1. Effectiveness of Machine Learning Models: The study evaluates fourteen different machine learning models, including linear, non-linear, probabilistic, tree-based, and neural network models, for performance prediction.
2. Impact of Hardware Type: The accuracy of performance predictions varies significantly between simulation-based hardware and physical systems.
3. Application Characteristics: The paper highlights that memory-bound applications tend to have higher prediction accuracy compared to compute-bound applications.
4. Dataset Diversity: The research utilizes a diverse set of datasets, including nine different databases from scientific applications executed on both simulated and physical systems.

3. Predicting Computer Performance Based on Hardware Configuration Using Multiple Neural Networks

a. Abstract of the research paper:

How can we accurately predict the performance of a Personal Computer (PC) configuration without time consuming simulation? In this work, we predict the performance of a computer hardware configuration using Multiple Neural Networks (MNN). We use Principal Component Analysis (PCA) during data preprocessing as guidance for model creation. The input data includes the internal component characteristics of a computer. A deep learning model is used to infer a benchmark score given a hardware configuration. The finished model takes input data such as Central Processing Unit (CPU) type, frequency, number of cores, memory size and speed, flash or disk architecture, network configuration and correlates it against the corresponding performance benchmark value and system response to a benchmark workload. We demonstrate the accuracy and effectiveness of the MNN and PCA machine learning models using the Standard

Performance Evaluation Corporation (SPEC) benchmarks (SPEC CPU2006 and SPEC CPU2017), and a set of approximately 50,000 commercial machines configurations. Our MNN model is able to achieve an average accuracy rate of 97.5% for all benchmarks. Our results provide both personal and enterprise users a tool that can accurately estimate system configuration performance without lengthy and resource intensive benchmarking sessions.

b. Inference drawn from the paper:

1. SPEC CPU Dataset: The data set in this study consists of approximately 50,000 records. Each suite has a subset of different benchmarks.
2. Feature selection and extraction: In PCA analysis, they identified the most relevant and independent features correlated with the desired outcomes, to select a subset of hardware characteristics that improve model accuracy.
3. Predictive Model: The feedforward ANN are multivariate statistical model used the hidden layer's parameters like the number of nodes, to achieve better performance.

4. Review On Text Classification Using Improved Deep Learning Models

a. Abstract of the research paper:

Modern deep learning models have exhibited better performance than conventional machine learning methods across a range of text classification tasks, encompassing natural language inference, sentiment analysis, news categorization, and question answering. The application of text classification in natural language processing extends to tasks like spam detection, news categorization, information retrieval, sentiment analysis, and intent assessment. Traditional text classifiers utilizing machine learning techniques encounter challenges such as sparse data, dimensionality issues, and limited generalization capabilities. In contrast, classifiers based on deep learning effectively tackle these limitations, eliminating the need for intricate feature extraction processes, and providing improved learning capabilities along with enhanced prediction accuracy. One notable instance of Convolutional neural networks are used in this way (CNNs). This paper delineates the text classification process, emphasizing the application of deep learning methodologies for text classification.

b. Inference drawn from the paper:

1. Feature Extraction: CNNs automatically extract relevant features from text data without the need for extensive manual feature engineering. They utilize convolutional layers to apply filters that can detect various patterns.
2. Local Connectivity: The convolutional layers in CNNs focus on local regions of the input data, allowing the model to learn local features effectively.
3. Multiple Filter Sizes: CNNs can employ multiple filters of different sizes to capture features at various levels of granularity.
4. Max Pooling: After the convolutional layers, CNNs typically use max pooling to reduce the dimensionality of the feature maps.

5. Text Document Classification Using Deep Learning Techniques:

a. Abstract of the research paper:

The advancement in technology has resulted in expansion in the volume of online text documents. It is interesting to note that in the previous two years alone, more electronic data has been created than ever before by the entire human species. As a result, it is now essential to accurately classify this data according to their content, which also helps in further processing and the extraction of valuable features. Text based document classification is one of the very important problems in Natural Language Processing (NLP). Manual document classification techniques rely heavily on human power to examine and label documents based on their content. Whereas, traditional Machine Learning (ML) based algorithms require manual feature extraction prior to classification which requires choosing the best algorithm to extract handcrafted features. Both these strategies are not only time-consuming but also prone to error, and require choosing the best available algorithms. On the other hand, Deep Learning (DL) based algorithms do not require human intervention as they perform deep feature extraction and classification automatically with much better performance than the traditional ML based frameworks. In this paper, we present a completely automated and robust document classification method to classify online digital documents using DL based methods i.e. BERT and RoBERTa. The proposed technique achieved highest accuracy of 98.9% and can be deployed to classify digital text documents with a high performance.

b. Inference drawn from the paper:

1. Problem: The massive growth in electronic textual data demands efficient classification techniques to categorize and process this information.
2. Solution: The paper proposes a deep learning-based approach using BERT and RoBERTa, which automatically performs feature extraction and classification, eliminating the need for manual intervention.
3. Dataset: The models were trained and evaluated on a dataset of 2,225 documents from the BBC News website, categorized into five classes (business, entertainment, politics, sports, and tech).
4. Performance: RoBERTa achieved an accuracy of 98.9%, while BERT achieved 97.7%, both outperforming traditional machine learning approaches.

6. Double-Channel Short Text Classification With Fused Enhanced Features

a. Abstract of the research paper:

The present paper proposes a double-channel short text classification model with fused enhanced features to tackle the issues of feature sparsity, incomplete feature extraction, and important information loss in short text classification tasks. Specifically, Bert and CNN are employed to extract the global and local features of short text, respectively. Moreover, an attention mechanism is incorporated to enhance channel features and mitigate the problem of feature sparsity in short text. To prevent the loss of important information, the SoftPool pooling mechanism is adopted instead of the traditional average pooling. Then, the global and local features are fused interactively to address the issue of incomplete feature extraction. Finally, the Softmax classifier is utilized to obtain the final results for short text classification. The experimental results indicate that the model proposed in this study achieved optimal classification performance across three short-text datasets.

b. **Inference drawn from the paper:**

1. Global and Local Feature Extraction: BERT is used for extracting global features, capturing the contextual semantics of short texts, while CNN extracts local features.
2. SoftPool Pooling Mechanism: Instead of traditional pooling, SoftPool is used to retain more significant information, improving the effectiveness of the extracted features.
3. Dual-Channel Fusion: The global and local features are combined interactively to create a comprehensive feature set, addressing incomplete feature extraction in short text classification.
4. Performance: The model, named MC-BCEN, is tested on three datasets (THUCNews, SMP2020, and a shopping review dataset), showing superior performance over single-channel models. It achieves the highest classification accuracy and F1 scores compared to other models.

7. A Text Classification Method with an Effective Feature Extraction Based on Category Analysis:

a. **Abstract of the research paper:**

Text classification refers to determine the class of an unknown text according to its content in the given classification system. In order to extract fewer features to express the information in the text as much as possible, the paper analysis the various features' statistical properties and to extract the global features according to Zipf's law; and then, based on the statistical analysis of the features' classified information, the efficient feature is extracted by computing the contribute of a feature; After that, the traditional TF-IDF formula is improved using category frequencies named by TF-IDF-CF for calculating the feature weight; Finally the text classification method is proposed. The experiment results illustrate that feature extraction methods proposed in the paper are effective and the formula TF-IDF-CF for calculating the feature weight has higher classification accuracy.

b. Inference drawn from the paper:

The paper presents a two-step feature extraction method:

1. Global Feature Extraction: Based on Zipf's Law, it removes overly frequent or rare terms (which tend to be noise) by setting thresholds.
2. Partial Feature Extraction: This method selects features that are more representative of specific categories by analyzing their category frequencies.
3. Modified TF-IDF: The authors improve the traditional TF-IDF weight formula by integrating category frequency, which reflects how important a feature is to a particular category. This adjustment enhances the discriminative power of features, leading to more accurate classifications.
4. Experimental Validation: The proposed method is tested on the standard Newsgroup dataset, and the results show that the new approach outperforms the traditional TF-IDF method, yielding higher precision, recall, and F-measure in text classification tasks.
5. The paper concludes that combining effective feature extraction and category analysis leads to better classification accuracy and efficiency, reducing dimensionality while preserving important textual information.

8. Research on Dual-channel Text Feature Extraction Method Based on Neural Network

a. Abstract of the research paper:

In the existing research on Chinese natural language processing text feature extraction has always been a core problem in this field, and the advantages and disadvantages of text feature extraction performance are important factors that directly affect the performance of text processing technology. This research offers a dual-channel feature extraction method based on neural networks in order to handle this challenge of text feature extraction from the perspectives of local information and context global information of text sentences. The text's local features are extracted using GCN, while the text's overall semantic features are extracted using BiLSTM, and the full connected neural network is used to fuse the local eigenvector with the global eigenvector to obtain the final features of the text. To confirm the efficiency of the suggested feature extraction technique, the suggested approach is used to classify news material in this research. The outcomes of the experiments demonstrate that the performance of the suggested method outperforms that of the GCN and BiLSTM feature extraction methods used independently. When compared to other feature extraction approaches, this method's experimental results are still better.

b. Inference drawn from the paper:

1. Dual-channel Feature Extraction:

   GCN is used to extract local features by constructing a graph where nodes represent words and edges represent the relationship between them, leveraging mutual information to weigh the connections.

   BiLSTM is employed to extract global features, capturing both forward and backward context in a text sequence, providing a comprehensive understanding of the text.
2. Fusion of Local and Global Features: The method fuses the local features extracted by GCN with the global features from BiLSTM using a fully connected neural network. This fusion ensures that both local word relationships and overall semantic context are considered in the final feature vector.
3. Experimental Validation: The model was tested on a Chinese news dataset, where it outperformed other feature extraction methods like GCN, BiLSTM, TextCNN, and TextRNN. The GCN+BiLSTM combination achieved the best results, showing superior performance in accuracy, precision, recall, and F1-score.
4. Advantages: This dual-channel approach addresses limitations in existing methods. By combining the strengths of both, it provides a more holistic view of the text, improving the accuracy of tasks like text classification.

2.2. Exiting Systems

Existing systems for building a budget gaming PC primarily revolve around online tools, platforms, and communities that provide users with resources to select compatible components, check performance benchmarks, and ensure cost-effectiveness. Here are some key systems:

1. PCPartPicker:

- **Overview**: PCPartPicker is one of the most popular platforms for building custom PCs, including budget gaming setups.
- **Features**: It allows users to choose components (CPU, GPU, motherboard, RAM, etc.), ensuring compatibility between parts. The platform also shows price comparisons from various retailers, user reviews, and build guides.
- **Strengths**: Real-time price tracking, user-generated builds, and a community-driven recommendation system.
- **Limitations**: Limited in-depth performance analytics or personalized gaming recommendations beyond hardware compatibility.

2. Newegg PC Builder:

- **Overview**: Newegg, an online electronics retailer, offers a PC Builder tool that helps users assemble a custom PC based on their budget and gaming needs.
- **Features**: It provides access to Newegg's inventory of components, detailed specifications, and user reviews. The tool also highlights discounts and promotions on individual parts.
- **Strengths**: Direct purchasing from the platform, sales-based pricing, and easy component swapping.
- **Limitations**: Lacks advanced filtering for performance-based recommendations and doesn't always focus on balancing budget vs. performance.

3. BuildMyPC:

- **Overview**: BuildMyPC is another web-based tool similar to PCPartPicker, designed to help users create their ideal gaming PC within a specified budget.
- **Features**: The platform focuses on pricing and compatibility, ensuring that selected parts fit within the desired price range and will work together.
- **Strengths**: Simple interface, with a focus on providing budget constraints upfront.
- **Limitations**: Less comprehensive than PCPartPicker in terms of user builds, performance analytics, and price comparison across multiple sources.

4. Reddit (r/buildapc):

- **Overview**: Reddit's "buildapc" subreddit is a large, community-driven platform where users seek advice on building gaming PCs, including budget setups.
- **Features**: Users post build recommendations and get advice on component selection, compatibility, and performance for specific games.
- **Strengths**: Community feedback, personalized advice, and access to real-world experiences.
- **Limitations**: Recommendations can be subjective, and it lacks structured tools for checking compatibility or tracking component prices.

5. Tom's Hardware PC Build Guides:

- **Overview**: Tom's Hardware offers curated build guides for different price points, including budget gaming PCs.
- **Features**: These guides provide detailed explanations of each component, why it was chosen, and how it performs for gaming.
- **Strengths**: Expert advice and up-to-date component recommendations, with detailed performance benchmarks.
- **Limitations**: Static recommendations that may become outdated as hardware and prices change; lacks dynamic build customization options.

6. CyberPowerPC and iBUYPOWER:

- **Overview**: These companies provide pre-built gaming PCs and custom PC-building services, where users can select components for custom builds.
- **Features**: The systems are built by professionals, and users can choose configurations from a wide range of budgets, with optional component upgrades.
- **Strengths**: Convenience for users who prefer pre-built systems, with a range of prices and options.
- **Limitations**: Less flexibility in part selection compared to building a PC yourself, and can sometimes be more expensive than DIY solutions.

While existing systems like PCPartPicker and Newegg's PC Builder focus on ease of part selection and compatibility checking, they often lack advanced performance metrics and personalized recommendations for gaming. Community-driven platforms like Reddit and Tom's Hardware offer valuable advice but can be inconsistent or outdated. Most existing systems are strong in ensuring hardware compatibility but could benefit from more advanced tools for optimizing performance-to-cost ratios, especially for gamers working within tight budgets

2.3. Lacuna in the existing systems

1. **Lack of Consolidated Information for System Requirements and Hardware Compatibility**:
   The current ecosystem is fragmented, with information about system requirements and hardware compatibility scattered across multiple websites, forums, and vendor platforms. Gamers often need to navigate through various sources—official game websites, hardware manufacturers, and third-party reviews—to gather the necessary details. This lack of a centralized repository makes it time-consuming and challenging to compare requirements across different games or identify hardware configurations that can support a wide range of titles. Without a unified platform, users may struggle to make informed decisions, leading to suboptimal gaming experiences or unnecessary hardware purchases.
2. **Difficulty in Staying Updated with the Latest Gaming Hardware Trends**:
   The gaming hardware landscape evolves rapidly, with frequent releases of new processors, graphics cards, memory modules, and other components. For users, especially those not deeply immersed in the tech world, keeping track of the latest trends, performance benchmarks, and compatibility updates is a daunting task. Many existing platforms either lack timely updates or do not provide insights into emerging technologies, making it difficult for gamers to stay current. As a result, users may end up purchasing outdated hardware or fail to take advantage of the latest advancements in gaming technology.
3. **Inadequate Tools for Performing In-Depth Compatibility Analysis**:
   Existing tools often provide limited functionality when it comes to analyzing hardware compatibility in-depth. While some websites offer basic comparison features, they rarely account for the complex interactions between different components, such as CPU-GPU bottlenecks, memory bandwidth limitations, or power supply constraints. Furthermore, they fail to consider how different games perform on various configurations under real-world conditions. This lack of detailed analysis tools forces users to rely on manual research or educated guesses, increasing the risk of compatibility issues, poor system performance, or unnecessary upgrades.

2.4. Comparison of existing systems and proposed area of work

Feature/Criteria	Existing Systems	Proposed System
Component Compatibility Check	Tools like **PCPartPicker**, **Newegg PC Builder**, and **BuildMyPC** provide robust compatibility checks between components.	The proposed system will similarly offer component compatibility checks but will incorporate real-time performance data from benchmarks and reviews.
User Personalization	Most existing systems provide static recommendations based on user input but lack in-depth personalization beyond price/budget.	The proposed system will offer **personalized recommendations** based on the user's budget, preferred games, and required performance levels, providing tailored setups with real-time data insights.
Feature Correlation	Existing tools focus on simple component matching (e.g., **PCPartPicker** checks if CPU is compatible with motherboard).	The proposed system will introduce a **feature correlation matrix** that identifies not just compatibility but also optimal component pairings for gaming performance (e.g., ideal CPU-GPU combinations).
Health Check/Optimization	No existing system offers a comprehensive health check for existing PCs.	The proposed system will include a **PC health check** feature to analyze the user's current system and recommend upgrades or optimizations based on performance bottlenecks.

Association Rule Mining	**No existing system** provides insights into frequently used or successful hardware combinations (e.g., CPU-GPU pairings).	The proposed system will use **association rule mining** to discover common and successful component pairings that yield optimal gaming performance within a given budget.
Driver Recommendations	No explicit feature in existing systems to suggest specific drivers based on the user's system configuration.	The proposed system will provide **driver recommendations** based on the user's OS and hardware configuration, improving system stability and gaming performance.
Ease of Use	Existing systems like **Newegg PC Builder** and **BuildMyPC** offer simple interfaces for users to create their builds.	The proposed system will offer an equally **user-friendly interface** with additional customization options for tailoring performance and budget recommendations.
Data Sources for Analysis	Primarily price data from retailers, static performance data from reviews, and product specs.	The proposed system will integrate data from **YouTube channels PC Build reviews**, hardware review sites, and **social sentiment analysis** from gaming communities, creating a more holistic recommendation system.

Table No. 1: Comparison of existing system and proposed system

2.5. Focus Area

The focus areas of the proposed system for building a budget gaming PC are centered around providing a personalized, data-driven approach to hardware recommendations, performance optimization, and system compatibility. Key focus areas include:

1. **Personalized Hardware Recommendations:**

   The system will focus on **personalizing gaming PC builds** based on the user's budget, preferred games, and performance expectations. It will provide tailored recommendations, ensuring the user gets the best possible gaming experience within their financial constraints.

2. **Component Compatibility and Optimization:**

   Ensuring **compatibility between hardware components** is a critical focus. The system will automatically check that all selected components (e.g., CPU, GPU, motherboard, RAM) work together without issues and suggest the most **optimal component pairings** for gaming performance.

3. **PC Health Check and Optimization Suggestions:**

   Another focus area is the ability to perform a **health check on existing gaming systems**, identifying hardware bottlenecks or underperforming components. The system will suggest upgrades or optimizations to boost overall system performance, ensuring users can extend the lifespan of their current setups.

4. **Driver and Software Recommendations:**

   The system will not only suggest hardware but also recommend **driver updates** based on the user's operating system and hardware configuration. This ensures users have the latest drivers for optimal performance and stability in gaming.

5. **Advanced Feature Correlation and Association Rule Mining:**

   The system will focus on developing a **feature correlation matrix** and use **association rule mining** to discover frequent patterns in hardware combinations. This will help identify the best hardware pairings that consistently deliver high performance for budget gaming setups.

6. **Trend-Based Pre-Built Configurations:**

   The system will focus on offering **trend-based pre-built configurations** that reflect current popular gaming setups. This allows users to quickly choose from **tested and proven configurations** that align with their budget and gaming requirements.

Summary of Focus Areas:

- **Personalization**: Tailored gaming PC builds based on budget and performance.
- **Health Check**: Evaluate existing systems and suggest upgrades or improvements.
- **Driver Recommendations**: Ensure optimal software support for gaming hardware.
- **Ease of Use**: Provide a user-friendly, accessible platform for all types of users.

By focusing on these areas, the proposed system will offer a comprehensive, data-driven, and personalized experience for building and optimizing budget gaming PCs.

Chapter 3: Requirements for the proposed system

3.1. Proposed System

The proposed system is a dynamic, personalized platform designed to help users build and optimize budget gaming PCs. It focuses on providing real-time recommendations by integrating data from various sources such as gaming platforms, hardware reviews, and performance benchmarks. The system ensures component compatibility and suggests optimal hardware configurations based on the user's budget, gaming preferences, and desired performance.

Key features include a PC health check to analyze the user's current system, driver recommendations based on their setup, and association rule mining to identify successful hardware pairings. Additionally, the system performs sentiment analysis on user reviews to offer insights into common hardware issues or advantages, and it provides pre-built configurations based on the latest hardware trends.

The user-friendly interface, combined with cross-platform accessibility, allows users to easily input their specifications and receive tailored recommendations, ensuring the best gaming performance within their budget.

3.2. Functional Requirements

1. **Data Extraction and Processing :**

   The system shall automatically scrape game and hardware data from Steam, Epic Games, and Games of Galaxy. The system shall retrieve system requirements for various games, including minimum and recommended hardware specifications. The system shall extract captions and descriptions from the scraped data for further analysis. It shall perform sentiment analysis on user reviews to gauge user satisfaction and identify key trends. It shall identify and catalog essential hardware features such as CPU, GPU, RAM, and storage requirements.

2. **Configuration Analysis:**

   The system shall analyze pre-built PC configurations available on the market. It shall identify and track trending configurations based on user preferences and performance metrics.

3. **Feature Correlation:**

   The system shall build a feature correlation matrix that illustrates relationships between different hardware components.It shall allow users to view how specific hardware combinations affect gaming performance.

4. **Compatibility Analysis:**

   The system shall utilize data from YouTube transcripts, GPU benchmarks, and hardware reviews to check compatibility between hardware components and games. It shall cross-reference the feature correlation matrix with the compatibility analysis results to ensure optimal hardware setups.

5. **Association Rule Mining:**

   The system shall implement association rule mining techniques to identify frequent patterns among hardware components.It shall provide users with recommended hardware combinations that have been successful in gaming setups.

6. **PC Health Check:**

   The system shall evaluate the health of the user's current PC setup based on performance metrics. It shall suggest optimizations and upgrades based on the health evaluation, focusing on improving gaming performance.

7. **Driver Recommendations:**

   The system shall provide links to the latest drivers for identified hardware components based on the user's operating system. It shall ensure that the driver recommendations are relevant to the user's current hardware configuration.

8. **Website Development:**

   The system shall create a static website that offers guidance on building optimal gaming setups.The website shall feature content on hardware compatibility, optimization tips, and driver updates.

3.3. Non-Functional Requirements

Here are the non-functional requirements for the project aimed at providing personalized gaming setup recommendations:

1. **Scalability**:

   The system shall be designed to accommodate an increasing volume of data and users, ensuring that performance remains consistent as more games and hardware configurations are added. The system shall allow for the integration of additional data sources in the future without requiring a complete overhaul.

2. **Reliability**:

   The system shall ensure high availability, with a target uptime of at least 99.5%. The system shall include error-handling mechanisms to manage failures gracefully and provide meaningful feedback to users.

3. **Usability**:

   The user interface shall be intuitive and easy to navigate, requiring minimal effort for users to find and utilize features. The system shall provide clear and concise instructions for users, especially during the input process for hardware specifications.

4. **Data Quality**:

   The system shall ensure that the data collected from external sources is accurate, up-to-date, and free from duplicates. The system shall regularly validate and clean the data to maintain its integrity.

5. **User Feedback**:

   The system shall include mechanisms for users to provide feedback on recommendations and functionality, facilitating continuous improvement. The system shall track user interactions to identify areas for enhancement and potential new features based on user needs.

These non-functional requirements outline essential quality attributes that the project must achieve to ensure a successful and satisfying user experience while maintaining robustness and adaptability

3.4. Constraints

Here are some potential constraints for the proposed system aimed at providing personalized gaming setup recommendations:

1. **Technical Constraints**:
    - **Data Source Limitations**: The system relies on the availability and accessibility of data from external sources like Steam, Epic Games, and Games of Galaxy. Changes to their APIs or scraping restrictions could impact data collection.
    - **Performance Limits**: The system must operate within the performance limits of the hosting infrastructure. Resource constraints may affect the speed and efficiency of data extraction and processing.
2. **Time Constraints**:
    - **Project Timeline**: The project must be completed within a specified timeframe, which may limit the depth of features implemented and the thoroughness of testing.
    - **Updates and Maintenance**: Regular updates to the game and hardware databases may be required, necessitating a commitment of time and resources to keep the system current.
3. **User Constraints**:
    - **User Hardware Limitations**: The recommendations provided may need to consider the user's current hardware limitations and preferences, which could limit the range of suggestions.

- **Diverse User Experience Levels**: The system must cater to users with varying levels of technical expertise, which can complicate interface design and feature accessibility.

4. **Dependency Constraints**:
    - **Third-party Libraries and Tools**: The system may depend on various third-party libraries and tools for functionalities like web scraping and sentiment analysis, which can introduce potential points of failure or compatibility issues.
    - **Internet Connectivity**: Users must have reliable internet access to utilize the system effectively, which may exclude some users in areas with poor connectivity.
5. **Data Quality Constraints**:
    - **Quality of Extracted Data**: The accuracy and completeness of the recommendations depend heavily on the quality of the scraped data. Inaccurate or incomplete data can lead to suboptimal recommendations.
    - **Latency in Data Updates**: The frequency of updates from external data sources may lead to a lag in the recommendations being current or reflective of the latest trends in hardware and games.
6. **User Interface Constraints**:
    - **Mobile Compatibility**: The design must consider mobile users, which may limit certain features or layouts that work better on desktop interfaces.
    - **Accessibility Requirements**: The system must adhere to accessibility guidelines to ensure it is usable for individuals with disabilities, which can impose additional design constraints.

3.5. Hardware & Software Requirements

1. **Hardware**: Servers for data storage and processing.
2. **Software:** Web scraping tools, NLP libraries, machine learning frameworks, web development tools.
3. **Tools:** Python, BeautifulSoup, Selenium, Power Automata, TensorFlow, Keras, PyTorch, HTML, CSS, JavaScript, Docker.

3.6. Techniques utilized for the proposed system

Here are the key techniques that can be utilized for the proposed system aimed at providing personalized gaming setup recommendations:

1. **Web Scraping**:
    - **Technique**: Use web scraping tools (e.g., Beautiful Soup, Scrapy) to extract data from websites such as Steam, Epic Games, and Games of Galaxy.
    - **Purpose**: Gather system requirements, game metadata, user reviews, and other relevant data for processing and analysis.
2. **Natural Language Processing (NLP)**:
    - **Technique**: Employ NLP techniques such as tokenization, named entity recognition (NER), and sentiment analysis using libraries like NLTK or SpaCy.
    - **Purpose**: Analyze user reviews to extract sentiment and identify key features and trends related to gaming experiences.
3. **Data Analysis**:
    - **Technique**: Use statistical analysis methods and tools like Pandas and NumPy to process and analyze the collected data.
    - **Purpose**: Extract insights from the data, such as common hardware configurations and performance metrics.
4. **Machine Learning**:
    - **Technique**: Implement machine learning algorithms (e.g., decision trees, random forests, or neural networks) for classification tasks using frameworks like Scikit-learn or TensorFlow.
    - **Purpose**: Predict compatibility and performance outcomes based on user hardware specifications and game requirements.
    - **Purpose**: Create a resource hub for users to access information on compatibility, optimization tips, and driver updates.

By leveraging these techniques, the proposed system can effectively collect, analyze, and present data to provide users with personalized gaming setup recommendations that ensure compatibility and optimize performance.

3.7. Tools utilized for the proposed system

Here’s a concise list of tools that can be utilized for the proposed system aimed at providing personalized gaming setup recommendations:

1. **Web Scraping Tools**:
    - **Beautiful Soup**: For parsing HTML and XML documents to extract game and hardware data.
    - **Selenium:** A powerful web automation tool that can simulate user interactions with websites, enabling dynamic web scraping of content that requires JavaScript execution.
    - **Power Automate:** A Microsoft tool designed for automating repetitive tasks across various applications, which can also be used for web scraping through its desktop flows to automate browser interactions.
2. **Natural Language Processing Libraries**:
    - **NLTK (Natural Language Toolkit)**: For text processing tasks, sentiment analysis, and linguistic feature extraction.
    - **SpaCy**: For advanced NLP tasks including tokenization and named entity recognition..

Chapter 4: Proposed Design

4.1. Block diagram

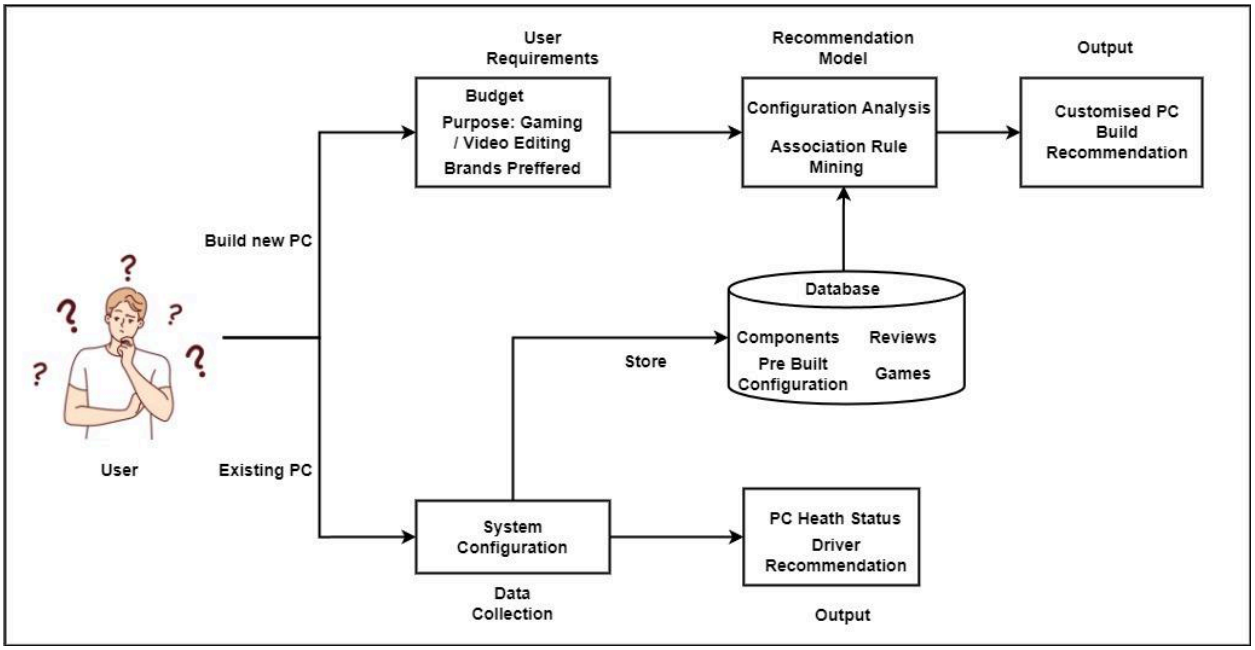

Fig 1: Block diagram

Summary:

1. **Build New PC:**

   If a user wants to build a new PC, he/she needs to enter the details like what brands does the user prefer buying from or if components are selected by user and other needs to be suggested, what would be the usage of the PC, which application would the user be using.

   The model has been trained on several datasets like reviews from which sentiment analysis has been performed, the pre-built configurations are analyzed which would help generate new ones, components data has been used for generating rules using association rule mining, the games define the system requirements.

   Thus the trained model would combine the user requirements and generate customized, compatible PC Build Configuration.

2. **Existing PC:**

   The user can get a PC Analysis done by just one click. Here the Batch file collects the system information data and generates a PC health status report.

   Additionally, the system would recommend the latest drivers available from the official website of the components manufacturer site.

4.2. Modular diagram

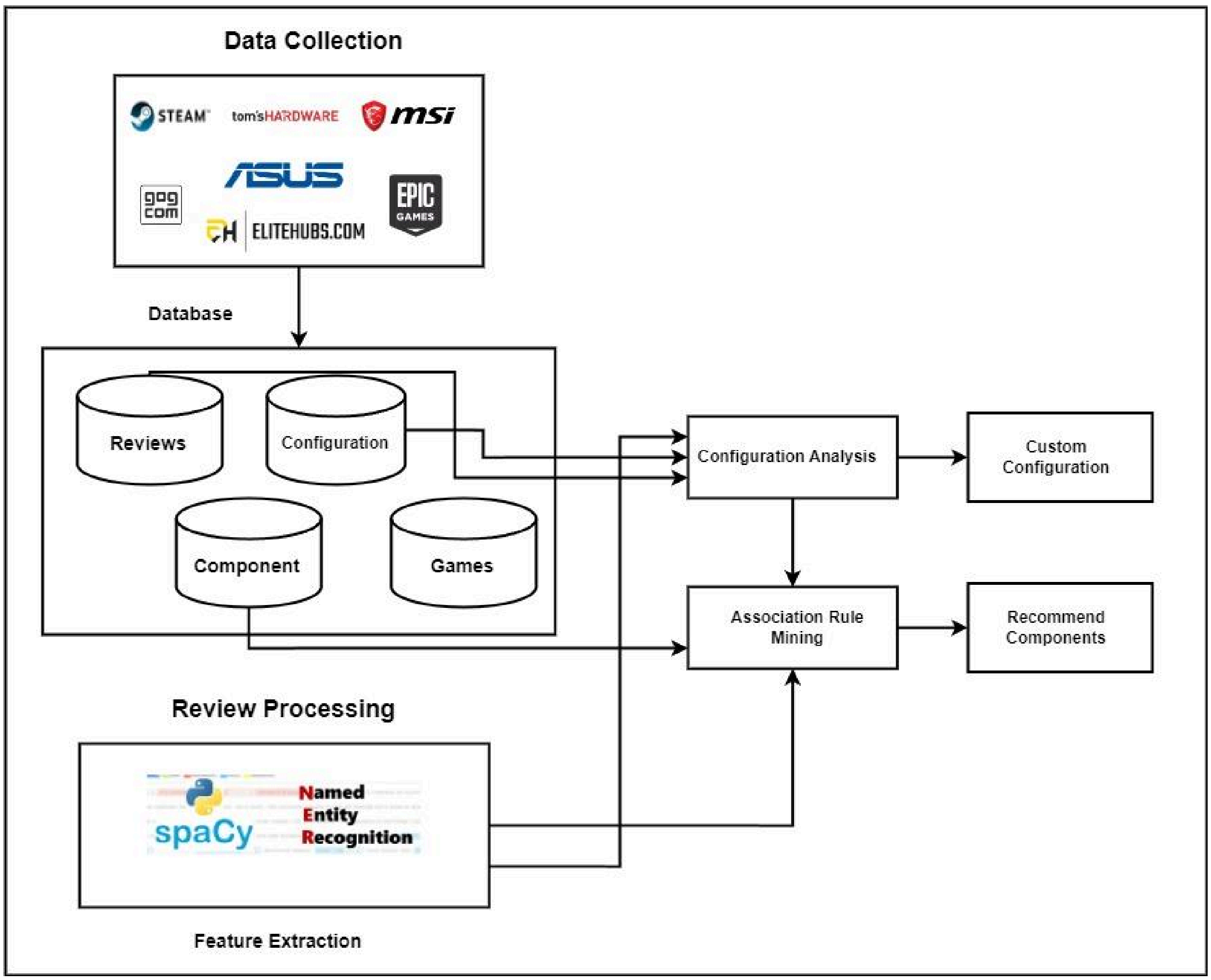

Fig 2: Modular Diagram

Components:

1. **Data Collection:**

   The data has been collected for various components: CPU, PSU, GPU, RAM, Motherboard, Monitor from various manufacturers like Asus, Asrock, MSI, Gigabyte, Cooler Master etc, the games data have been collected from steam application, and other sites, review are collected from Tom's Hardware, Jazy two cents youtuber's channels (some sites as listed in references). The data has been processed and stored in a structured manner.

2. **Review Processing:**

   The reviews data needs to be processed wherein the features like products name, company' name, the dimensions are to be extracted from the text and used for further processing. Several models like BERT, RoBERT, Spacy etc are used and the most accurate one is chosen.

3. **Configuration Analysis:**

   This uses majorly two databases, the pre-built configuration and reviews one, it maps the review to the configuration which will be further used to predict whether the configuration is compatible or not.

4. **Association Rule Mining:**

   The association rule mining is used to generate the buying patterns for the compatible products which would be recommended to the user. This is particularly useful for users with a limited or mid-range budget.

4.3. Algorithms utilized in the proposed systems

1. BERT (Bidirectional Encoder Representations from Transformers):

- **Overview**: BERT is a **transformer-based** machine learning model developed by Google for natural language processing (NLP) tasks. It is pre-trained on a large corpus of text and designed to understand the context of words in a sentence by processing text bidirectionally.
- **Key Features**:
  - **Bidirectional Training**: Unlike traditional models that read text from left-to-right or right-to-left, BERT reads the entire sentence in both directions simultaneously, helping it understand context better.
  - **Pre-training and Fine-tuning**: BERT is pre-trained on tasks like masked language modeling and next sentence prediction, and can then be fine-tuned for specific tasks such as text classification, named entity recognition (NER), and question answering.
  - **Applications**: BERT is widely used for tasks like sentiment analysis, question answering, and NER due to its deep contextual understanding of language.
- **Advantages**: BERT's bidirectional approach allows it to capture richer, more nuanced word meanings based on context.

2. RoBERTa (A Robustly Optimized BERT Pretraining Approach):

- **Overview**: RoBERTa is an enhanced version of BERT, developed by Facebook, which improves upon BERT by optimizing its training strategies. It stands for "Robustly Optimized BERT Pre-training Approach."
- **Key Differences from BERT**:
  - **Larger Training Data**: RoBERTa is trained on more data compared to BERT, including larger and more diverse datasets.
  - **Longer Training**: RoBERTa extends the pre-training duration for better results, removing the next-sentence prediction task and focusing solely on masked language modeling.

    - **Dynamic Masking**: RoBERTa applies dynamic masking during training, meaning that different words are masked across different epochs.
- **Applications**: Like BERT, RoBERTa is used for various NLP tasks, such as text classification, NER, and question answering, but it generally performs better due to these optimizations.
- **Advantages**: RoBERTa typically outperforms BERT on most NLP benchmarks due to more efficient training techniques and optimized use of resources.

3. SpaCy:

- **Overview**: SpaCy is a popular open-source library for advanced natural language processing tasks in Python. It is designed to be **fast and efficient**, with pre-trained models for different languages and tasks like named entity recognition, part-of-speech tagging, dependency parsing, and more.
- **Key Features**:
    - **Speed**: SpaCy is optimized for real-time applications, offering fast performance for processing large volumes of text.
    - **Pre-trained Models**: It provides out-of-the-box pre-trained models for different languages and offers the ability to fine-tune them for custom NLP tasks.
    - **Entity Recognition**: SpaCy has an efficient named entity recognition (NER) system that can identify entities such as names, dates, and locations within a text.
    - **Applications**: It is used in a wide range of NLP tasks, including entity recognition, sentiment analysis, text classification, and machine translation.
- **Advantages**: SpaCy is known for its easy-to-use interface, integration with deep learning libraries, and high performance, making it ideal for large-scale applications.

4. Stanford NER (Named Entity Recognizer):

- **Overview**: Stanford NER is a **Conditional Random Fields (CRF)-based model** developed by the Stanford NLP Group for identifying named entities like names, dates, organizations, and locations in text. It is part of the Stanford CoreNLP toolkit.

- **Key Features**:
  - **Entity Detection**: Stanford NER focuses on recognizing and classifying named entities in text, which is essential for tasks like information extraction, question answering, and summarization.
  - **CRF-Based Approach**: Stanford NER relies on Conditional Random Fields, a statistical modeling method, to predict and classify named entities. It uses both local word features and global sequence features for better accuracy.
  - **Customization**: It can be fine-tuned on custom datasets for specific domains (e.g., biomedical or legal text) to recognize custom entities.
- **Applications**: Widely used in academic and enterprise applications for tasks such as text mining, information extraction, and automatic content generation.
- **Advantages**: Stanford NER is a mature, accurate tool for NER, especially for research purposes. However, it tends to be slower compared to transformer-based models like BERT.

Summary of Differences:

- **BERT and RoBERTa** are transformer-based models, with RoBERTa being a more optimized and robust version of BERT, both excelling in tasks that require understanding context in large amounts of text.
- **SpaCy** is a fast and practical library for various NLP tasks, including NER, with a strong emphasis on performance and ease of use.
- **Stanford NER** is a traditional, CRF-based model for named entity recognition, known for its accuracy but not as fast or adaptable as newer transformer-based models.

These tools are widely used in natural language processing tasks, with BERT and RoBERTa leading in modern, context-aware applications, while SpaCy and Stanford NER are popular in both industry and academic environments for various NLP use cases.

5. Linear Regression:

Linear Regression is a statistical method used to model the linear relationship between a dependent variable (target) and one or more independent variables (predictors). It aims to predict the value of the dependent variable based on the independent variable(s) by fitting a linear equation to the observed data.

Key Features:

- Simple Linear Regression: Involves modeling the relationship between a single independent variable and a dependent variable using a straight line. The equation is expressed as: y=mx+by = mx + by=mx+b

  Where m is the slope and b is the intercept.
- Multiple Linear Regression: Extends simple linear regression to include two or more independent variables, modeling their combined effect on the dependent variable.
- Model Fitting: The linear regression algorithm fits the best possible line through the data points by minimizing the difference between the predicted and actual values, usually using a method called least squares.

Advantages:

- Simplicity: Linear regression is easy to implement and interpret, making it a great starting point for regression analysis.
- Efficiency: Works well for datasets with a clear linear relationship between variables.
- Scalability: Scalable to larger datasets, especially when using multiple variables (multiple regression).

Limitations:

- Assumes Linearity: It assumes a linear relationship between variables, which may not always hold true in real-world data.
- Sensitive to Outliers: Outliers can significantly affect the performance and accuracy of the model.

Chapter 5: Proposed Results and Discussions

5.1 Implementation:

The user system data is collected using the following commands and stored in a file. Also there are commands that generate the whole system report and give the health status of each component.

The screenshot attached for the same are as follows:

1. **User System data:**

```
F:\Projects\gaming-python\scripts>powershell -Command "Get-WmiObject
rer, Model, SystemType"


Manufacturer : Gigabyte Technology Co., Ltd.
Model        : B460MDS3HV2
SystemType   : x64-based PC
```

Fig 3: Motherboard Details

```
F:\Projects\gaming-python\scripts>powershell -Command "Get-WmiObject
rer, PartNumber, SerialNumber, Speed, Capacity, DeviceLocator, BankLa


Manufacturer  : Kingston
PartNumber    : KHX2666C16/8G
SerialNumber  : BAA7761A
Speed         : 2666
Capacity      : 8589934592
DeviceLocator : ChannelA-DIMM0
BankLabel     : BANK 0

Manufacturer  : Kingston
PartNumber    : KF2666C16D4/8G
SerialNumber  : 8006A163
Speed         : 2666
Capacity      : 8589934592
DeviceLocator : ChannelB-DIMM0
BankLabel     : BANK 2
```

```
F:\Projects\gaming-python\scripts>powershell -Command "Get-WmiObject
apacity, MemoryDevices"


MaxCapacity   : 67108864
MemoryDevices : 4
```

Fig 4: RAM Details

2. **Health Status Of PC:**

```
Disk Information:
Caption             Size           Status
ST1000DM010-2EP102  1000202273280  OK
ZEB-SD13            128034708480   OK
```

Fig 5: Disk Details

```
==============================
Memory Information
==============================

Gathering information about physical memory...
BankLabel  Capacity    DeviceLocator   Manufacturer  MemoryType  PartNumber      SerialNumber  Speed  Tag

BANK 0     8589934592  ChannelA-DIMM0  Kingston      0           KHX2666C16/8G   BAA7761A      2666   Physical Memory 0

BANK 2     8589934592  ChannelB-DIMM0  Kingston      0           KF2666C16D4/8G  8006A163      2666   Physical Memory 2



Gathering system memory status...
Total Physical Memory:     16,305 MB
Available Physical Memory: 4,550 MB

Gathering virtual memory details...
Virtual Memory: Max Size:  36,310 MB
Virtual Memory: Available: 14,058 MB
Virtual Memory: In Use:    22,252 MB

==============================
Script completed.
==============================
Press any key to continue . . .
```

```
=========================================================
System Information Report
=========================================================

CPU Information:
Manufacturer  MaxClockSpeed  Name                                      NumberOfCores  NumberOfLogicalProcessors
GenuineIntel  2904           Intel(R) Core(TM) i5-10400F CPU @ 2.90GHz  6              12
```

Fig 6: RAM & CPU Details

```
=====================================================
System Information Report
=====================================================

GPU Information:
GPU detected: NVIDIA GeForce GTX 1050 Ti
=====================================================
NVIDIA
Running NVIDIA Query...
temperature.gpu, fan.speed [%], power.draw [W], clocks.current.graphics [MHz]
39, 45 %, [N/A], 1290 MHz
=====================================================
Press any key to continue . . .
```

Fig 7: GPU Details

3. **Feature Extraction:**

   Features like the product name being discussed in the text, the manufacturer name, series and its other details are extracted as follows:

... Lenovos Lunar Lake-powered X1 Carbon Aura Edition PRODUCT weighs just 2.16 pounds Prior ThinkPad X1 Carbons SERIES weighed about 0.3 pounds more At IFA 2024 Lenovo is bringing Intel ORG s Lunar Lake SERIES CPUs to its flagship ThinkPad SERIES while chopping off a lot of the laptops weight. The ThinkPad X1 Carbon Gen 13 Aura Edition PRODUCT tips the scales at just 2.16 pounds which is quite a bit lighter than its 2.42-pound predecessor. Available in November for a starting price of 1999 this super-slim 14-inch laptop promises an AI-enhanced user experience long battery life and plenty of performance and flexibility. Designed to compete with the best ultrabooks the X1 Carbon Gen 13 PRODUCT has the word Aura in its name because Lenovo ORG is starting its Aura program in which the company worked with Intel ORG to add a few extra utilities that enhance the user experience. Along with the X1 Carbon Gen 13 PRODUCT Lenovos new Yoga Slim 7i PRODUCT is also an Aura Edition SERIES laptop. Both Auras have three major features In addition to these Aura AI features the X1 Carbon Gen 13 PRODUCT will have the Copilot PC experiences that Microsoft has made available on Qualcomm Snapdragon CPU -powered laptops. These include live captions Windows Studio Effects MISC for your webcam and Cocreator a text-to-image feature of Windows Paint MISC . The X1 Carbon PRODUCT comes with one of Intels new Intel Core Ultra Series 2 SERIES processors otherwise known as Lunar Lake SERIES . It will be available with up to a Core Ultra 9 SERIES CPU no word on the bottom end but we expect Core Ultra 7 SERIES and Ultra 5 SERIES SKUs with up to 32GB SIZE of RAM RAM and up to a 2TB SIZE SSD STORAGE . According to the

Fig 8: Terms recognised from article

4. **Association Rule Mining:**

   This displays the probability of the products being bought together:

```
antecedents,consequents,antecedent support,consequent support,support,confidence,lift,leverage,conviction,zhangs_metric
frozenset({'Processor_Intel Core i5 12400'}),frozenset({'Motherboard_MSI Pro H610M-E DDR4'}),0.14285714285714285,0.2857142857142857,0.1428
frozenset({'Motherboard_MSI Pro H610M-E DDR4'}),frozenset({'Processor_Intel Core i5 12400'}),0.2857142857142857,0.14285714285714285,0.1428
frozenset({'RAM_Adata XPG Gammix D30 DDR4 16GB'}),frozenset({'Processor_Intel Core i5 12400'}),0.2857142857142857,0.14285714285714285,0.14
frozenset({'Processor_Intel Core i5 12400'}),frozenset({'RAM_Adata XPG Gammix D30 DDR4 16GB'}),0.14285714285714285,0.2857142857142857,0.14
frozenset({'Graphics Card_Integrated GPU'}),frozenset({'Processor_Intel Core i5 12400'}),0.14285714285714285,0.14285714285714285,0.142857
frozenset({'Processor_Intel Core i5 12400'}),frozenset({'Graphics Card_Integrated GPU'}),0.14285714285714285,0.14285714285714285,0.142857
```

Fig 9: Rules extracted

5. **Pre-trained Models Performance Metrics:**

The confusion matrix for these pretrained models is as follows for feature extraction, the accuracy is quite low therefore manual annotation of the data is required:

```
True Positives: 7 - {(0, 5, 'ORG'), (647, 652, 'ORG'), (851, 854, 'ORG'), (980,
False Negatives (missed): 1151 - {(2925, 2932, 'ORG'), (467, 477, 'ORG'), (1293,
False Positives (wrong predictions): 29 - {(238, 247, 'NORP'), (1436, 1438, 'ORG
Accuracy: 0.005862646566164154
Precision: 0.19444444444444445
Recall: 0.006044905008635579
```

Fig 10: Confusion matrix for Spacy model

```
True Positives: 12 - {('SEGA', 'ORG'), ('SMIC', 'ORG'), ('Reuters', 'ORG'), ('Acer',
False Negatives (missed): 143 - {('Tesla', 'LOC'), ('Dominic', 'ORG'), ('Lunar', 'OR
False Positives (wrong predictions): 246 - {('Be Quiet Silent Base 802', 'CABINET'),
Accuracy: 0.029925187032418952
Precision: 0.046511627906976744
Recall: 0.07741935483870968
```

Fig 11: Confusion matrix for Stanford NER

```
True Positives: 55 - {('SMIC', 'ORG'), ('Cortex AI', 'MISC'), ('Gigabyte Technologies
False Negatives (missed): 472 - {('Ram', 'MISC'), ('New', 'ORG'), ('G300', 'MISC'), (
False Positives (wrong predictions): 527 - {('Nvidia Blackwell', 'SERIES'), ('Iomega',
Accuracy: 0.05218216318785579
Precision: 0.09450171821305842
Recall: 0.10436432637571158
```

Fig 12: Confusion matrix for BERT

```
True Positives: 77 - {('SMIC', 'ORG'), ('Iomega', 'ORG'), ('VideoCardz', 'MISC'), ('Be Quiet', 'ORG'),
False Negatives (missed): 445 - {('G300', 'MISC'), ('Tom Hardware', 'ORG'), ('Quad Royal', 'MISC'), (
False Positives (wrong predictions): 505 - {('Nvidia Blackwell', 'SERIES'), ('2 GHz', 'METRIC'), ('Lu
Accuracy: 0.07497565725413827
Precision: 0.1323024054982818
Recall: 0.1475095785440613
```

Fig 13: Confusion matrix for RoBERTa

Chapter 6: Plan of action for the next semester

6.1. Work done till date:

1. **Data Extraction:** CPU, GPU, RAM, Power Supply Units, Motherboard, Monitors, Games, Pre Built Configuration sites, Youtube reviews, Text reviews.
2. **Data Preprocessing:** Data was collected from various sources hence transformed into one defined universal format, handling of missing values, feature selection was performed.
3. **Feature Extraction:** Features like company name, product name, its details like dimension are extracted from the textual reviews articles.

6.2. Plan of action for project II:

1. Build a custom recommendation model for custom model configuration generation.
2. Build a website wherein users can interact with the whole system.

Chapter 7: Conclusion

The proposed system for building a budget gaming PC provides a comprehensive, personalized approach to optimizing gaming performance within a user's financial constraints. By integrating real-time data from various platforms, such as benchmarks, hardware reviews, and gaming community feedback, the system offers tailored recommendations that go beyond simple component compatibility. It introduces advanced features like **performance-based analysis**, **sentiment-driven insights**, and **association rule mining**, ensuring users receive configurations that meet their gaming requirements while optimizing their budget.

The system's **PC health check** and **driver recommendation** features extend its utility beyond initial builds, helping users maintain and optimize their systems over time. With its user-friendly interface and focus on accessibility, the platform offers a seamless experience for both novice and experienced gamers, guiding them through the complexities of building and maintaining a gaming PC. Ultimately, this project addresses key limitations in existing systems and delivers a more dynamic, intelligent solution for budget gaming PC enthusiasts.

Chapter 8: References

1. https://store.steampowered.com/
2. https://store.epicgames.com/en-US/
3. https://www.gog.com/en
4. https://www.youtube.com/c/jayztwocents
5. https://www.youtube.com/channel/UCXuqSBlHAE6Xw-yeJA0Tunw
6. https://www.tomshardware.com/archive
7. https://www.primeabgb.com/

Chapter 9. Appendix

a. List of Figures

b. List of tables

c. Review Sheets:

1. **Review 1:**

Industry/Inhouse: **Project Evaluation Sheet 2024-25** (45) **Class: D17 A**

Title of Project(Group no): Gamespec : Advisor (45)

Group Members: Harsh Dadlani (09) Manan D., Ayush Rakesh Gena (18) Ayush, Meet Kewalramani (28) Meet, Anisha Sunil Rohra (52) Anisha

	Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg & Mgmt principles	Life-long learning	Professional Skills	Innovative Approach	Total Marks
	(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(5)	(5)	(50)
Review of Project Stage 1	04	04	04	03	04	02	02	02	02	02	02	02	04	04	41

Comments:

Priyanka Shinde
Name & Signature Reviewer1

	Engineering Concepts & Knowledge	Interpretation of Problem & Analysis	Design / Prototype	Interpretation of Data & Dataset	Modern Tool Usage	Societal Benefit, Safety Consideration	Environment Friendly	Ethics	Team work	Presentation Skills	Applied Engg & Mgmt principles	Life-long learning	Professional Skills	Innovative Approach	Total Marks
	(5)	(5)	(5)	(3)	(5)	(2)	(2)	(2)	(2)	(3)	(3)	(3)	(5)	(5)	(50)
Review of Project Stage 1	04	04	04	03	04	02	02	02	02	02	02	02	05	05	42

Comments:

Date: 23rd August, 2024

Richard Joel Sujata Ph
Name & Signature Reviewer2

2. Review 2:

Industry/Inhouse: Inhouse **Project Evaluation Sheet 2024-25** **Class: D17** A

Title of Project(Group no): GameSpec Advisor (45)

Group Members: Ayush Jeevra (18), Anisha Rohra (52), Neet Kewalramani (28), Manan Dadlani (9)

	Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (3)	Applied Engg &Mgmt principles (3)	Life - long learning (3)	Professional Skills (5)	Innovative Approach (5)	**Total Marks (50)**
Review of Project Stage 1	04	04	04	03	04	02	02	01	01	03	02	02	04	04	40

Comments:

Name & Signature Reviewer1

	Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (3)	Applied Engg &Mgmt principles (3)	Life - long learning (3)	Professional Skills (5)	Innovative Approach (5)	**Total Marks (50)**
Review of Project Stage 1	04	04	04	03	04	02	02	01	01	03	02	02	04	04	40

Comments:

Date: 26th September, 2024

Name & Signature Reviewer2