SWINBURNE UNIVERSITY OF TECHNOLOGY

HO CHI MINH CAMPUS

COS30018 – Intelligent System

Group 8: VRP_Vehicle Routing System

Instructor: Aiden NGUYEN

Students Names:

Nguyen Van Quoc (104198640)

Mai Hoang Dai Vy (104993201)

Tran Hung Quoc Tuan (105000908)

Nguyen Ha Minh Chau (104663478)

HO CHI MINH CITY – March 2025

I. Introduction

The Vehicle Routing Problem (VRP) represents a critical optimization challenge in modern logistics, demanding intelligent solutions to efficiently manage delivery operations. Our project develops an advanced routing system that addresses the complex task of assigning delivery routes while maximizing operational efficiency. At its core, the system employs a multi-agent approach featuring a Master Routing Agent (MRA) and multiple Delivery Agents (DAs), designed to optimize route assignments by considering critical constraints such as vehicle capacity, maximum travel distances, and delivery priorities. The primary objective is to create a flexible routing solution that can handle real-world logistical challenges, with a specific focus on maximizing the number of items delivered while minimizing total travel distance. By implementing sophisticated search and optimization techniques, the system goes beyond traditional routing methods. It introduces innovative capabilities, including dynamic route optimization, adaptive handling of variable vehicle capacities, and an interactive user interface for comprehensive system configuration. Through this approach, the project demonstrates the powerful potential of optimization algorithms (both GA and OR-Tools) in solving complex transportation management problems, offering a glimpse into the future of intelligent logistics systems.

II. Overall system architecture

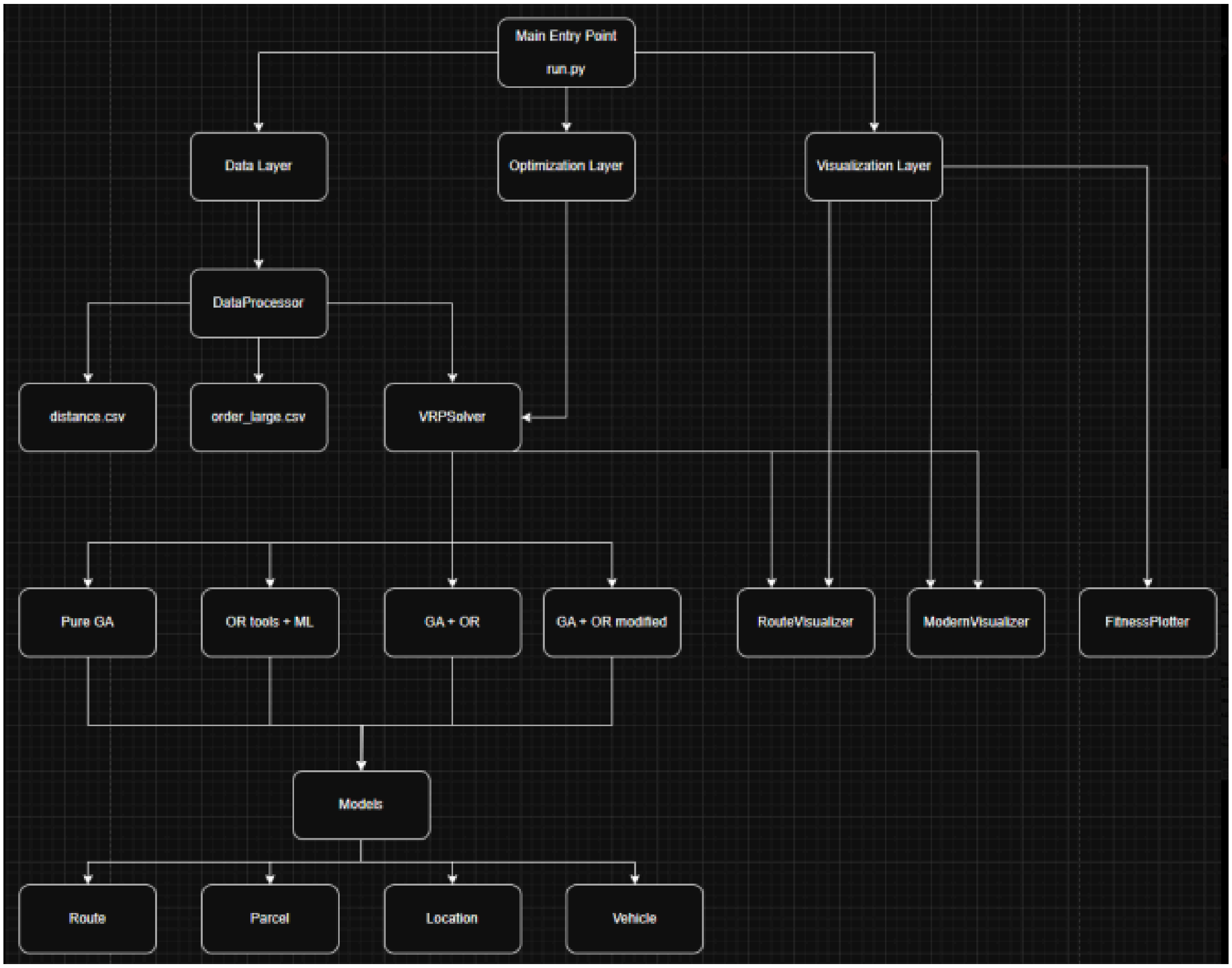


Figure 1.1 Overall System Architecture diagram

The diagram illustrates the system architecture of the Vehicle Routing System (VRP) project, outlining its layered structure and component interactions. The architecture is organized into three primary layers: the Data Layer, the Optimization Layer, and the Visualization Layer. The process begins at the run.py entry point, initiating the application's execution.

The Data Layer is responsible for managing input data, utilizing files such as distance.csv and order_large.csv, and employs a Data Processor to handle data extraction and preparation. The Optimization Layer contains the core VRP solving logic, using the VRPSolver to implement various optimization techniques, including Pure Genetic Algorithm (GA), OR-Tools with Machine Learning (ML), a hybrid GA + OR approach, and a modified version of the GA + OR hybrid. Finally, the Visualization Layer focuses on presenting the optimized routes and related information through components like RouteVisualizer, Modern Visualizer, and FitnessPlotter.

The system also employs several models, including Route, Parcel, Location, and Vehicle, representing the key entities within the VRP domain.

III. Implemented interaction protocols

Our Vehicle Routing Problem (VRP) solution implements sophisticated interaction protocols to deliver seamless communication between system components and optimization methods. These protocols point out how different system elements exchange information, coordinate decision-making, and collaborate to produce optimized routing solutions.

Message-Based Communication System

The protocols component includes three main communication methods:

1. Message Protocol

- Creates standard message formats that all system components can understand
- Ensures data is consistent and correctly formatted
- Handles errors when messages aren't properly formatted
- Allows components to send and receive information without waiting for responses

2. Agent Protocol

- Controls how the Master Routing Agent and Delivery Agents talk to each other
- Defines what each agent can do and how they should behave
- Sets up rules for assigning delivery routes to agents
- Manages the lifecycle of agents from start to finish

3. Message Queue

- Works like a mail sorting system for messages between components
- Makes sure messages reach their destination even when the system is busy
- Helps the system handle many operations at once

How Models Communicate with Other Components

Our model's components (Location, Parcel, Route, Vehicle) communicate to other parts of the system through:

1. Location and Parcel Data Protocol

- Provides simple ways to ask for location and delivery information
- Checks that location coordinates and parcel details are valid
- Helps verify that delivery constraints can be met

2. Route and Vehicle Interaction Protocol

- Defines how optimization algorithms can access and change routes
- Checks that vehicles don't exceed capacity or distance limits
- Helps compare different possible routes
- Converts route data between different optimization methods

Optimization Solution Selection Protocol

The system implements a flexible solution selection protocol that allows users to choose between multiple optimization approaches:

1. Pure Genetic Algorithm Protocol

- Evolution-based optimization with population management
- Communicates solution fitness metrics to the visualization layer
- Implements generational progression protocols with **selection**, **crossover**, and **mutation** operators

2. OR-Tools with ML Enhancement Protocol

- Communication between Google's OR-Tools solver and the ML model
- The ML component (Random Forest Regression) receives feature data (parcels, weights, and distances)
- Predicted route costs are fed back to the OR-Tools solver to enhance optimization

3. GA + OR-Tools Hybrid Protocol

- Sequential communication protocol where OR-Tools initializes solutions
- Solutions are transformed into a format compatible with GA evolution
- The feedback loop enables iterative refinement between the two methods

4. GA + OR-Tools with Modified Fitness Protocol

- Enhanced communication protocol with pattern analysis
- Adapted fitness calculations based on route characteristics
- Bidirectional information exchange for weighted optimization

Data Flow Communication Protocol

1. Data Processing Protocol

- Data_processor components extract and validate data from CSV sources
- Error-handling protocols manage data inconsistencies

2. Model-Algorithm Interaction Protocol

- Uniform interface for all optimization methods to interact with model objects
- Route, Parcel, Location, and Vehicle models expose standardized query methods

- Constraint validation protocols ensure solution feasibility

Hybrid Approach Coordination Protocol

1. GA + OR Integration Protocol

- OR-Tools solutions serve as seed data for the initial GA population
- Standardized conversion ensures compatibility between different solution representations
- Performance metrics are shared to guide the evolutionary process

2. Pattern Extraction Communication Protocol

- Pattern analysis components identify routing patterns from historical solutions
- Extracted patterns inform the weighted fitness calculations
- Adaptive influence mechanisms adjust optimization parameters based on performance feedback

Visualization of Communication Protocol

1. Route Visualization Protocol

- Standardized route data formatting for visual representation
- Update mechanisms for dynamic route rendering
- Interactive selection and filtering capabilities, allow users to select and filter routes

2. Performance Metrics Communication

- Integration with the UTILS Performance component
- Real-time metrics updates for fitness plotting
- Shows comparisons between different optimization methods

IV. Implemented optimization techniques

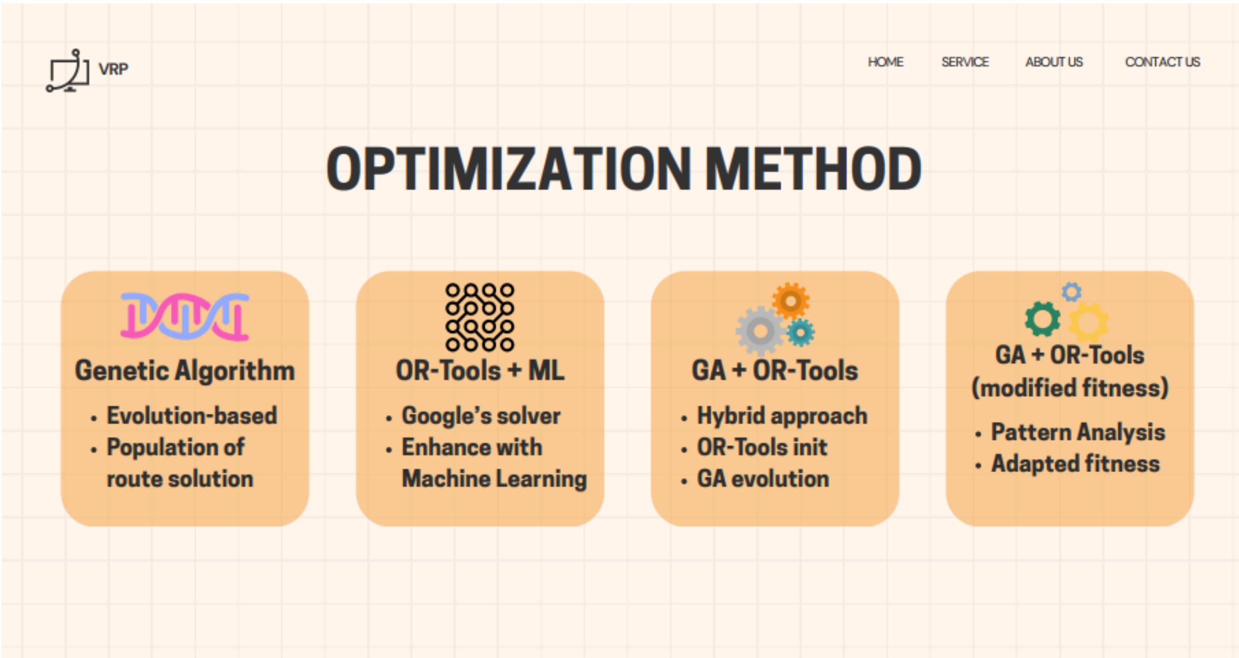


Figure 4.1 Optimization Method

- **Genetic Algorithm (GA)**: This is described as an evolution-based optimization method that works with a population of route solutions. The GA implementation involves several steps:

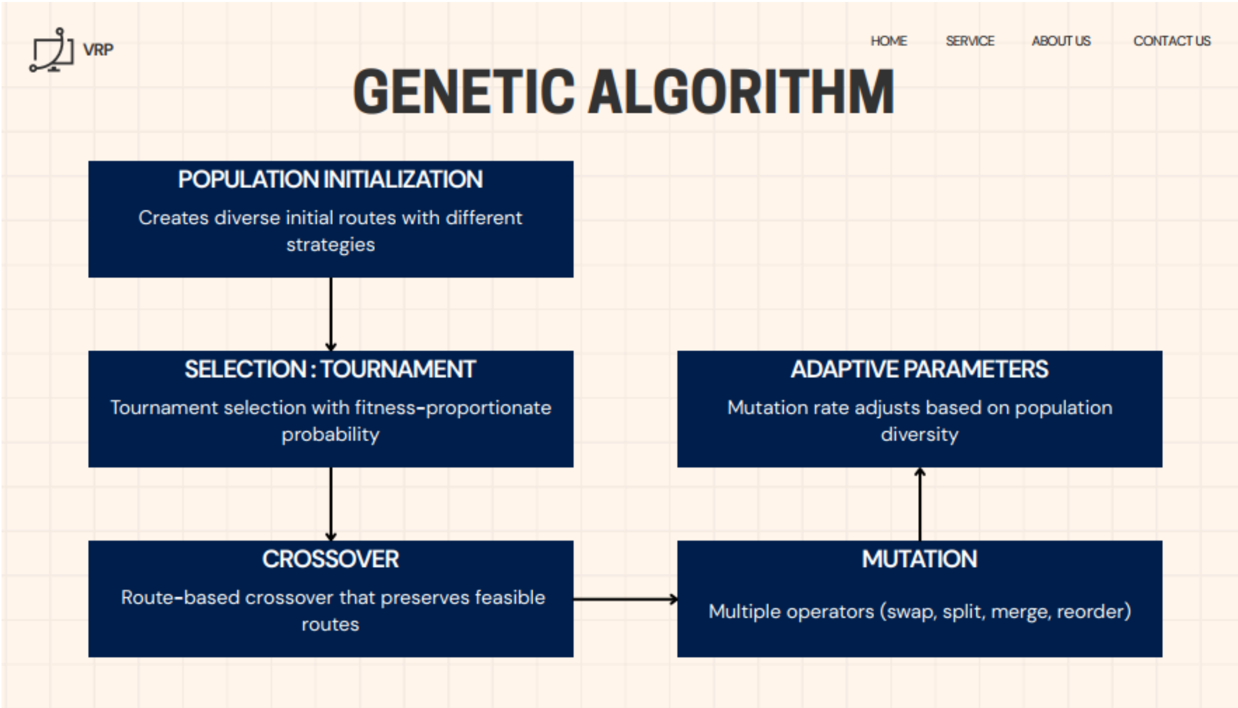


Figure 4.2 GA Workflow

1. **Population Initialization:** Creating diverse initial routes using different strategies.
2. **Selection:** Employing tournament selection with fitness-proportionate probability to choose routes for reproduction.
3. **Crossover:** Using route-based crossover that preserves feasible routes to create new solutions.
4. **Mutation:** Applying multiple operators (swap, split, merge, reorder) and adjusting the mutation rate based on population diversity. The source notes that a Pure GA approach is easily adaptable to new constraints and is highly customizable, with no dependency on external solvers.

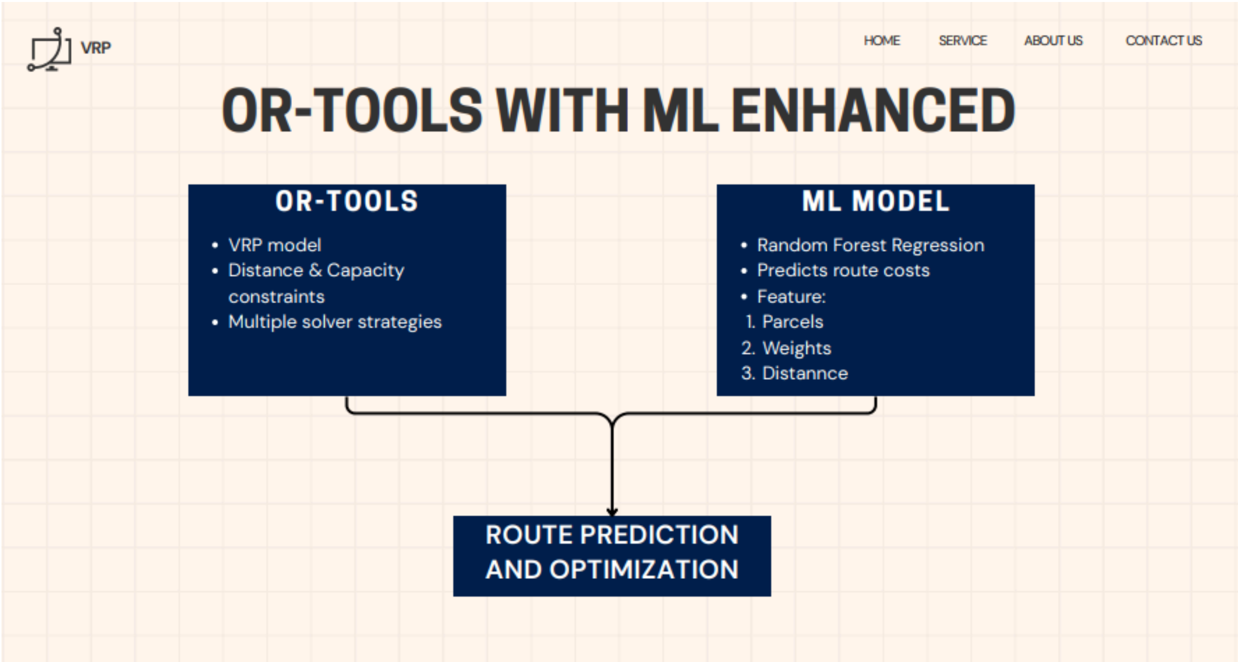


Figure 4.3 OR-Tools and ML combination

- **OR-Tools + ML:** This refers to Google's solver for optimization problems. The system utilizes OR-Tools by itself and also in conjunction with other methods. The OR-Tools model considers distance and capacity constraints and offers multiple solver strategies. On the other hand, a Machine Learning model, specifically Random Forest Regression, is used to predict route costs based on key features such as the number of parcels, their weights, and the total distance traveled.

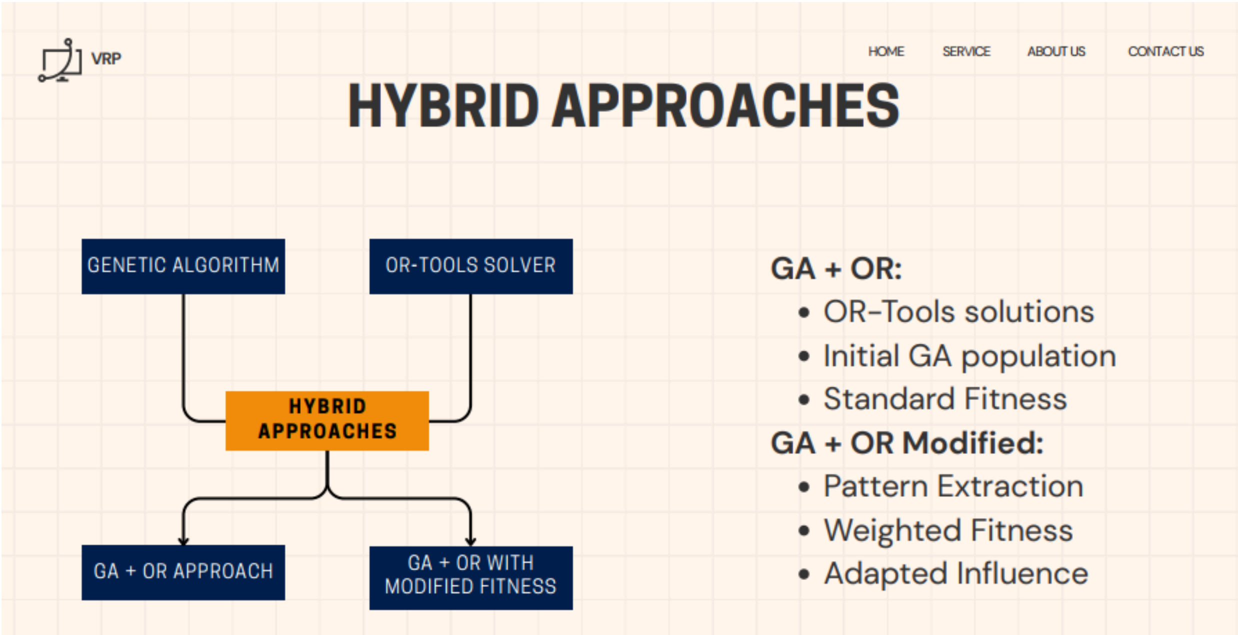
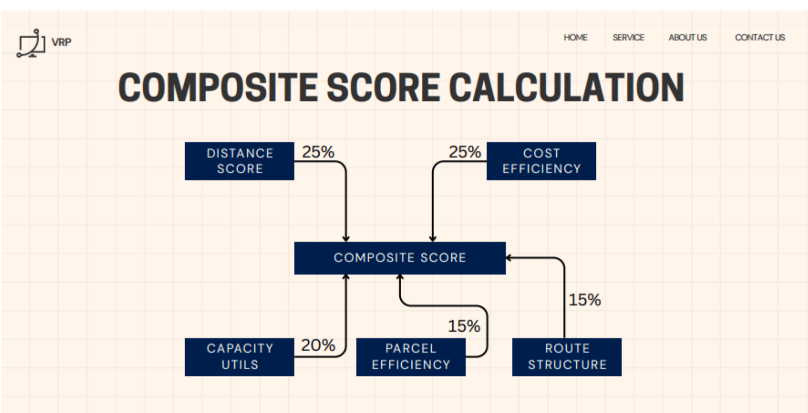


Figure 4.4 Hybrid Approaches

- **GA + OR:** This represents a hybrid approach where OR-Tools solutions are used to generate the initial GA population and then a standard fitness calculation is applied within the GA framework.
- **GA + OR with modified fitness:** This is another hybrid approach that builds upon the standard GA + OR method by incorporating pattern extraction to apply a weighted fitness calculation and achieve an adapted influence on the optimization process. This modified approach showed improvements in capacity utilization and better route structure scores according to the key findings.



The Composite Score is derived from five key metrics, each contributing a specific percentage weight to the final score:

Distance Score (25%) evaluates the total distance traveled by vehicles. Lower distances generally indicate better efficiency.

Figure 4.5 Composite Score Calculation

Cost Efficiency (25%) measures the cost-effectiveness of the route, considering fuel consumption, driver expenses, and other operational costs.

Capacity Utilization (20%) assesses how well the vehicle's load capacity is utilized, ensuring that trucks are neither underloaded nor overloaded.

Parcel Efficiency (15%) looks at how effectively parcels are packed and distributed to maximize delivery efficiency.

Route Structure (15%) examines the overall organization of the delivery route, ensuring logical and optimized routing to minimize unnecessary detours.

The composite score is calculated as a weighted average of five individual scores:

Distance Score: This score is weighted at 25%. An example formula provided is max(0, 1 - (avg_distance_per_parcel / (total_parcels * 100))).

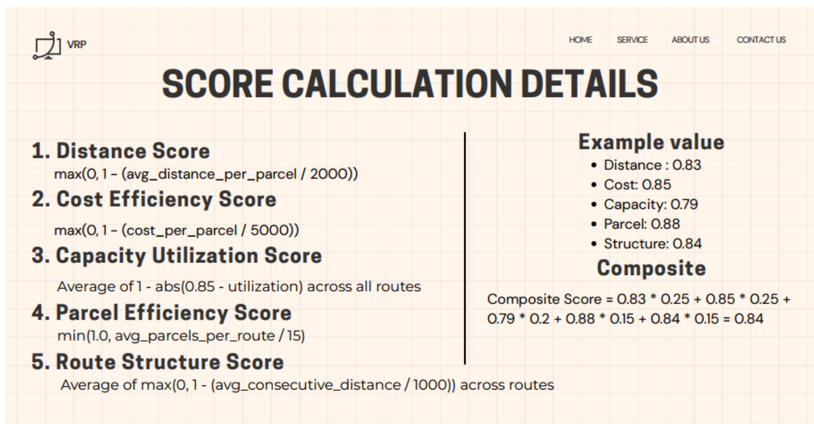


Figure 4.6 Score Calculation Details

Cost Efficiency Score: This score is also weighted at 25%. An example formula is max(0, 1 - (cost_per_parcel / (total_distance * 2.5))). Capacity Utilization Score: This score has a weight of 20%. The calculation is described as the "Average of 1 - abs(0.85 - utilization) across all routes".

Parcel Efficiency Score: This score is weighted at 15%. The formula is min(1.0, avg_parcels_per_route / 15). An example value given is 0.88. This score considers the average number of parcels delivered per route, potentially aiming for an average of 15.

Route Structure Score: This score also has a weight of 15%. The calculation is described as the "Average of max(0, 1 - (avg_consecutive_distance / 1000)) across routes". An example value given is 0.84. This score likely evaluates the structure and efficiency of the routes themselves, penalizing long consecutive distances.

V. Scenarios/examples to demonstrate how the system works

Route visualization is a crucial aspect, providing a clear representation of the optimized routes, aiding in solution comprehension, issue identification, and effective communication of results

The Pure Genetic Algorithm (GA), which employs evolutionary principles to generate diverse initial routes and iteratively improve them through selection, crossover, and mutation; OR-Tools + ML, which leverages Google's OR-Tools solver enhanced with machine learning for route prediction; and hybrid approaches, including GA + OR-Tools, combining the strengths of both techniques, and GA + OR-Tools with a modified fitness function, where pattern extraction and weighted fitness are used to adapt the influence of different factors on the optimization.

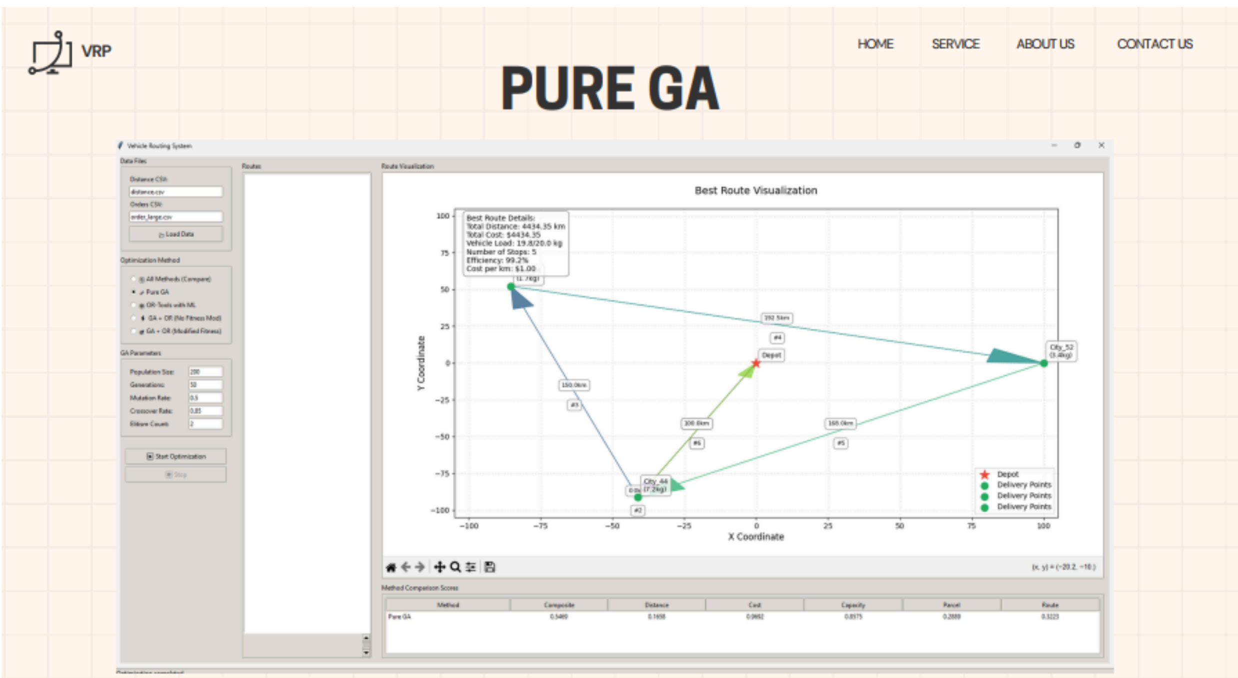


Figure 5.1 Pure GA

This is the application of a Genetic Algorithm (GA) for optimizing vehicle routes within a Vehicle Routing System (VRP). The interface consists of multiple components, including a distance data input section, GA optimization parameters. Users can load distance data and configure key GA settings such as population size, mutation rate, crossover rate, and elitism rate, which influence the efficiency of the optimization process. The best route visualization graphically represents the optimized delivery path, showing key locations such as the depot, delivery points, and the connections between them, while also displaying critical details such as total distance traveled, route cost, number of generations used, and cost per kilometer. At the bottom, a comparison table evaluates the GA's performance based on factors like composite score, distance efficiency, cost per kilometer, capacity utilization, and overall route efficiency.

This visualization highlights how a pure GA-based approach evolves route solutions over multiple generations, selecting and refining the best ones based on fitness scores. By leveraging this optimization technique, the system aims to generate efficient and cost-effective delivery routes, reducing operational expenses and improving logistics planning.

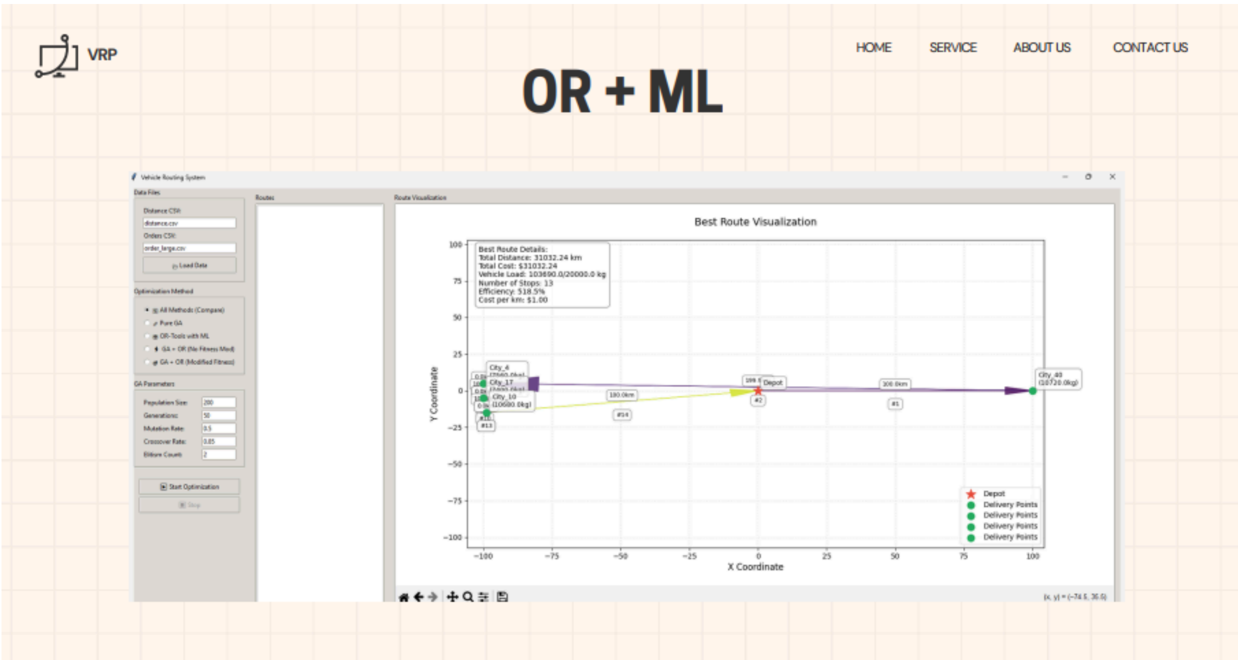


Figure 4.2 OR + ML

The OR + ML demonstrates the application of OR-Tools enhanced with Machine Learning for optimizing vehicle routes within a Vehicle Routing System (VRP). The interface displays a best route visualization, showing key locations such as the depot, delivery points, and the connections between them. The best route visualization graphically represents the optimized delivery path, along with critical details such as total distance traveled, route cost, vehicle load, number of stops, efficiency, and cost per kilometer. The graph plots delivery locations using X and Y coordinates, with the depot marked in red and delivery points in green. Different colored lines likely indicate different routes or segments. The side panel provides access to different optimization methods and parameters. This visualization highlights how the OR-Tools + ML approach generates an optimized route solution, focusing on efficiency and cost-effectiveness.

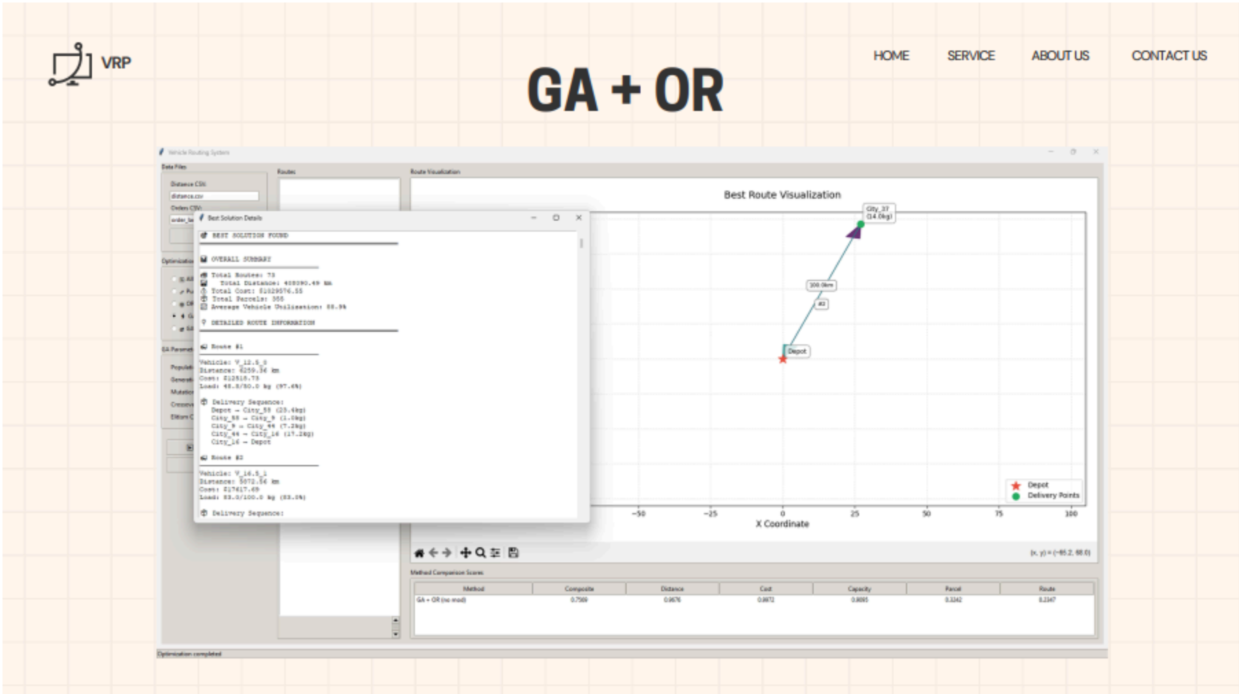


Figure 4.3 GA + OR

The GA + OR approach combines Genetic Algorithms (GA) with Operations Research (OR) techniques to optimize the Vehicle Routing Problem (VRP). In the first image, this method produces a relatively simple route, utilizing a single vehicle and minimizing travel distance. The visualization shows a direct delivery path with fewer stops, suggesting a focus on cost-effective and time-efficient routing. The best solution details indicate the total distance traveled, overall cost, time required, and vehicle capacity utilization. This approach is useful for scenarios where delivery points are limited and a straightforward optimization is needed. However, the single-vehicle approach might struggle with larger, more complex routing problems, where multiple deliveries need to be distributed efficiently.

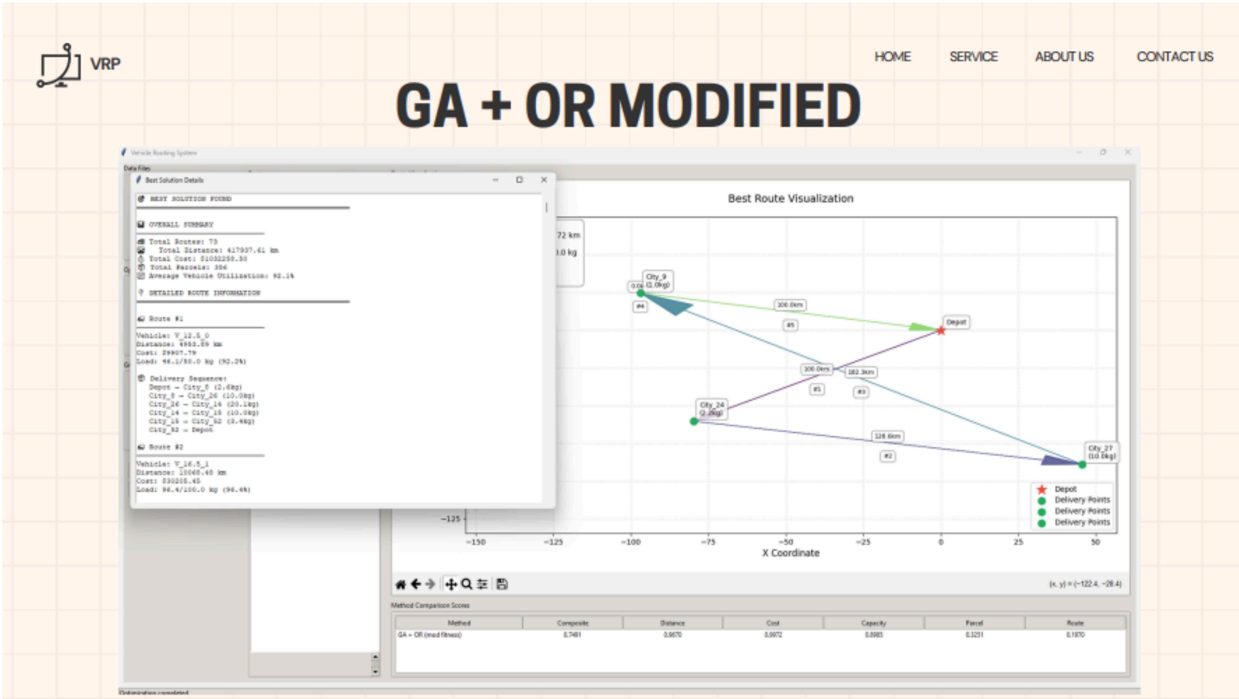


Figure 4.4 GA + OR Modified

The GA + OR Modified approach builds upon the original by introducing enhancements that improve efficiency. In the second image, this version of the algorithm optimizes routes using multiple vehicles, allowing for a more distributed workload across different delivery points. The route visualization is significantly more complex, reflecting the improved utilization of multiple routes and better capacity allocation. The modifications likely include advanced mutation, crossover strategies, or dynamic adjustments that reduce the total travel distance and improve cost efficiency. This approach is better suited for large-scale logistics operations, where multiple deliveries must be coordinated in an optimized manner. Compared to the standard GA + OR, the modified version achieves higher efficiency, better resource utilization, and a more effective distribution of deliveries, making it ideal for complex routing problems.

VI. Some critical analysis of the implementation

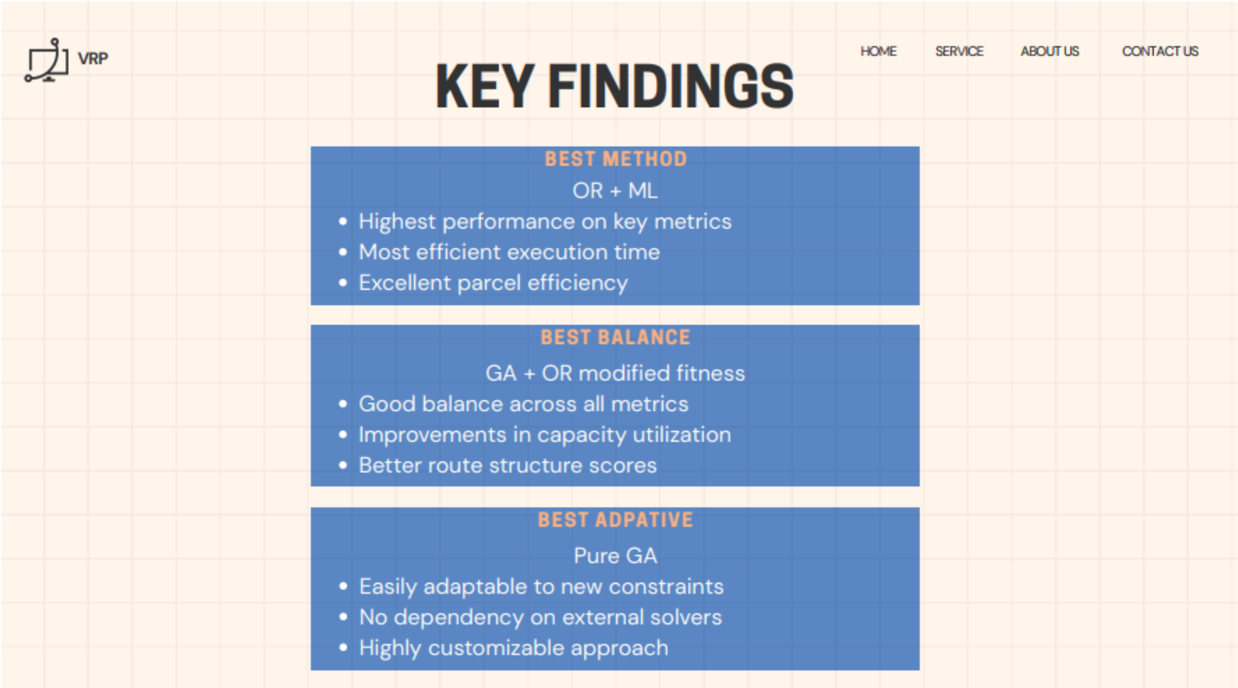


Figure 6.1 Key findings

The Best Method is identified as a combination of Operations Research (OR) and Machine Learning (ML), which provides the highest performance on key metrics, ensures the most efficient execution time, and offers excellent parcel efficiency.

The Best Balance approach involves Genetic Algorithms (GA) combined with OR and modified fitness functions, which results in a well-balanced performance across all metrics, significant improvements in capacity utilization, and better route structure scores.

Lastly, the Best Adaptive method is based on Pure GA, which stands out due to its ability to easily adapt to new constraints, lack of dependency on external solvers, and highly customizable approach.

In summary, this approach enhances OR-Tools with Machine Learning for route prediction, specifically using a Random Forest Regression model that predicts route costs based on features like parcels, weights, and distance. Key findings indicate that OR + ML demonstrated the highest performance on key metrics, the most efficient execution time, and excellent parcel efficiency

VII. Summary/Conclusion.

The OR-Tools with ML enhancement demonstrated the highest performance on key metrics, exhibited the most efficient execution time, and achieved excellent parcel efficiency1. The GA + OR with modified fitness approach presented a good balance across all metrics, showing improvements in capacity utilization and better route structure scores1. The Pure GA approach was noted for its easy adaptability to new constraints and its independence from external solvers, making it a highly customizable option

In summary, the VRP project has explored and evaluated different optimization strategies, identifying the strengths and weaknesses of each. The findings highlight the potential of hybrid approaches, especially OR-Tools enhanced with machine learning, for achieving high performance. Future work will focus on refining these optimization techniques, integrating real-world factors, and developing the necessary infrastructure for practical implementation