UJEAS

UNIZIK JOURNAL OF ENGINEERING AND APPLIED SCIENCES

UNIZIK Journal of Engineering and Applied Sciences 5(3), September (2025), 2816-2850
Journal homepage: https://journals.unizik.edu.ng/index.php/ujeas
PRINT ISSN: 2992-4383 || ONLINE ISSN: 2992-4391

Advanced Techniques for Fuel Blend Optimization Using Machine Learning, Thermodynamic Modeling and Experimental Validation Methods.

Sylvester Chukwutem Onwusa[1], Okotubu, Johnbull Oyonru[2], Okoye, Peter Izuoba[3] and Uyeri, Oghenerobo Cyril[4]

[1]Department of Mechanical Engineering, Delta State University of Science and Technology, Ozoro, Nigeria

[2]Department of Technical and Vocational Education, University of Delta, Agbor.

[3]Department of Technology and Vocational Education Nnamdi Azikiwe University, Awka, Nigeria

[4]Department of Mechanical Engineering, Delta State University of Science and Technology, Ozoro, Nigeria

*Corresponding Author's Email: onwusachukwutemsylvester@gmail.com

Abstract

The growing demand for cleaner and more efficient energy sources in internal combustion engines necessitates the development of optimized fuel blends. Conventional diesel, despite its high energy density, suffers from suboptimal combustion efficiency and elevated pollutant emissions. This study, titled Advanced Techniques for Fuel Blend Optimization (FBO) Using Machine Learning (ML) Thermodynamic Modeling (TDM) and Experimental Validation Methods (EVMs), aims to enhance engine performance and reduce emissions through the integration of alternative fuel blends (AFBs) and advanced predictive techniques. The primary objective is to identify optimal fuel formulations that outperform conventional diesel in thermal efficiency and environmental impact while leveraging modern machine learning (ML) and thermodynamic tools for performance forecasting and analysis. The methodology involves experimental testing of multiple diesel-diethyl ether (DEE) fuel blends, thermodynamic assessments including exergy and entropy analyses, and the application of three ML models Random Forest (RF), Extreme Gradient Boosting (XGBoost) and Artificial Neural Networks (ANN) to predict engine parameters such as Brake Thermal Efficiency (BTE), NO_x and CO emissions. Results show that Blend B4 (70% diesel + 30% DEE) achieved the highest BTE (34.0%), lowest BSFC (230 g/kWh), and significantly reduced emissions. XGBoost outperformed other ML models with R^2 values above 0.90 and lowest prediction errors (MAPE < 3%). The study concludes that oxygenated, high-cetane blends like B4 offer superior performance and environmental benefits. Furthermore, ML-based predictive models, particularly XGBoost, are reliable tools for real-time engine optimization. It is recommended that future research explore broader fuel types and integrate ML with real-time control systems for smart combustion management

Keywords: Advanced techniques, Fuel Blend Optimization, Machine Learning, Thermodynamic Modeling and Experimental Validation Methods**.**

1.0. Introduction

The escalating global energy demand, driven by rapid urbanization, industrial development and population growth, has intensified pressure on existing energy infrastructures, particularly within the transportation sector. Fossil fuels such as gasoline and diesel continue to dominate due to their high energy density, established supply chains and

compatibility with internal combustion engine (ICE) technologies (IEA, 2022; Demirbas, 2011). However, reliance on these fuels contributes substantially to greenhouse gas (GHG) emissions, especially carbon dioxide (CO_2), a leading driver of climate change. In addition to GHG emissions, fossil fuel combustion produces harmful pollutants such as nitrogen oxides (NO_x), hydrocarbons (HCs) and particulate matter (PM), which have been directly linked to deteriorating air quality and public health challenges including respiratory and cardiovascular diseases (Zhao et al., 2020; Reşitoğlu et al., 2015; Karavalakis et al., 2011).

Beyond their environmental and health impacts, fossil fuel dependency introduces significant economic and geopolitical risks, particularly through price fluctuations, supply insecurity, and reliance on finite reserves. These challenges have heightened global interest in developing sustainable alternatives that can ensure both energy security and environmental protection (Khan et al., 2020; Panwar et al., 2011). Among the strategies proposed, alternative fuel blends (AFBs) have attracted significant attention due to their potential to reduce life-cycle carbon intensity while maintaining compatibility with existing ICE platforms. These blends incorporate a range of renewable and synthetic options, including biofuels (such as biodiesel and bioethanol), alcohols (such as methanol and butanol), ethers (such as diethyl ether), and synthetic hydrocarbons derived from renewable feedstocks (Knothe, 2010; Hoekman et al., 2012; Demirbas, 2009). Importantly, they can often be integrated into current transportation infrastructures with minimal engine modifications, offering a practical pathway toward decarbonization (Lapuerta et al., 2008; Agarwal et al., 2013).
Despite these advantages, optimizing AFBs presents considerable challenges due to the complex interplay of their physicochemical properties. Parameters such as cetane number, oxygen content, volatility, viscosity and heating value exert strong influence on key combustion processes, including ignition delay, flame propagation and heat release rate. These combustion dynamics ultimately determine fuel efficiency, engine torque, and emission characteristics (Kumar et al., 2013; Labeckas & Slavinskas, 2006). For instance, biodiesel, with its high oxygen content, tends to reduce PM formation but often contributes to elevated NO_x emissions under certain operating conditions (Knothe, 2010; Lin et al., 2009). Alcohol-based fuels can stabilize combustion and improve thermal efficiency, but their low boiling points and hydrophilic nature introduce risks of volatility, corrosion and material incompatibility (Mustafa et al., 2012). The inherently nonlinear and multidimensional nature of these interactions complicates both predictive modelling and blend formulation (Rakopoulos et al., 2010; Agarwal et al., 2013).

Traditional research has largely depended on empirical testing and simplified thermodynamic models to investigate these trade-offs. While such studies, including those by Rakopoulos et al. (2010) and Onwusa et al. (2025), have provided valuable insights into biodiesel and related blends, they are resource-intensive and often lack scalability. Moreover, thermodynamic models grounded in the First and Second Laws of Thermodynamics remain essential but are frequently constrained by ideal assumptions that fail to capture transient engine behaviours and compositional variability in real-world applications (Heywood, 2018; Turns, 2013).

Conventional predictive approaches in fuel engine studies empirical regressions, design-of-experiments (DoE) response surfaces and stand-alone computational fluid dynamics (CFD) struggle with (i) the strongly nonlinear, multicollinear relationships among fuel physicochemical properties and engine responses; (ii) data sparsely across operating regimes and blend ratios and (iii) the need for mechanistic consistency when extrapolating beyond the test matrix. Simple regressions and low-order response surfaces assume near-linearity and weak interactions, which rarely hold when properties such as oxygen content, cetane number, volatility and viscosity jointly influence ignition delay, local equivalence ratio, heat-release phasing and post-oxidation chemistry (Kumar et al., 2013; Labeckas & Slavinskas, 2006; Rakopoulos et al., 2010). Meanwhile, high-fidelity CFD with detailed kinetics can be prohibitively computationally expensive for routine optimization over large blend/operation design spaces, and results are sensitive to sub-model choices and boundary condition uncertainty (Heywood, 2018; Turns, 2013). These limitations hinder scalability, complicate uncertainty quantification and can yield poor generalization to new fuels or duty cycles.

Machine learning (ML) addresses these deficits by learning complex, high-dimensional and nonlinear mappings directly from data, enabling accurate prediction of brake thermal efficiency (BTE), NO_x, CO, HC and PM across diverse AFBs and loads (Roy et al., 2021; Zhang et al., 2022; Sahin et al., 2022). Ensemble methods (RF, XGBoost) and deep neural networks (ANNs/DLNs) capture higher-order interactions (e.g., oxygen content × rail pressure × EGR rate) that elude low-order models, while built-in cross-validation and out-of-bag estimates support robust generalization (Chen & Guestrin, 2016; Sahin et al., 2022). Crucially, modern ML workflows can incorporate model interpretability and uncertainty tools permutation importance, partial dependence, SHAP value analysis, bootstrapped prediction intervals to expose which fuel properties and operating variables dominantly drive emissions and efficiency, countering the criticism that ML is a black box (Chia et al., 2020; Khan et al., 2020; Roy et al., 2021). Compared to

DoE/response surfaces, ML yields lower predictive error and better extrapolation control when appropriately regularized and validated on heterogeneous datasets (Zhang et al., 2022; Sahin et al., 2022).

Within this growing body of ML-driven work, diethyl ether (DEE) has emerged as a recurring test fuel due to its high oxygen content, volatility and excellent ignition quality, making it a common additive or blend component in biodiesel and alcohol-based studies. Prior ML-based investigations incorporating DEE have successfully modelled its effects on ignition delay, combustion phasing, and regulated emissions, underscoring both its potential and its relevance as a benchmark fuel in predictive analytics (Roy et al., 2021; Zhang et al., 2022; Sahin et al., 2022). However, these studies typically treat DEE in isolation or within limited binary blends, applying purely data-driven models without embedding thermodynamic consistency. This gap highlights the critical need for hybridized frameworks that move beyond isolated DEE optimization to generalizable, physically informed strategies applicable across diverse AFB systems.

Thermodynamic modelling (TDM) complements ML by enforcing first-principle consistency. First-Law energy balances, availability exergy analyses and equilibrium/finite-rate combustion calculations impose hard physical constraints on feasible temperature, pressure, species and efficiency outcomes (Moran et al., 2015; Turns, 2013; Heywood, 2018). Such constraints are indispensable when data are sparse or biased toward specific regimes, preventing ML from learning spurious correlations and ensuring predictions respect conservation and second-law limits. In practice, TDM supplies physically meaningful features/targets (e.g., net heat-release, specific exergy destruction, equilibrium adiabatic flame temperature) that improve learnability and reduce variance in ML models and it provides sanity checks for extrapolations (e.g., bounding NO_x formation via temperature–residence-time envelopes consistent with Zeldovich pathways; Turns, 2013).

The hybrid ML–TDM paradigm is therefore not merely advantageous but necessary. When ML is constrained or informed by thermodynamics through physics-guided feature engineering, penalty terms that enforce energy closure, or rejection of predictions violating equilibrium bounds the result is a mechanistically interpretable surrogate that retains ML's predictive power while avoiding nonphysical artifacts (Khan et al., 2020; Chia et al., 2020). This is especially critical for fuel blend optimization, where exploring large, continuous composition spaces (e.g., biodiesel–HVO–alcohol–ether mixtures) would be untenable using CFD alone and unreliable using unconstrained regressions. Embedding TDM in ML also enables data-efficient learning: small, carefully designed experiments calibrated with thermodynamic metrics can train models that generalize across engines and duty cycles, accelerating screening and down-selection before expensive engine/vehicle trials (Moran et al., 2015; Heywood, 2018; Roy et al., 2021).

What distinguishes the present study from earlier work is its advanced integration of model selection, tuning and validation protocols within the hybrid ML–TDM framework**.** Unlike prior DEE-focused or purely data-driven approaches, this study explicitly justifies the choice of Random Forest, ANN, and XGBoost based on their complementary strengths in capturing nonlinear interactions, conducting feature-level interpretation and maintaining robustness across heterogeneous AFB datasets. Hyper parameters are systematically optimized using grid search and Bayesian strategies to minimize over fitting, while k-fold cross-validation and out-of-sample validation ensure reproducibility and generalization. Furthermore, model interpretability (through SHAP, partial dependence) and uncertainty quantification (through bootstrap intervals) are embedded in the workflow, providing actionable insights rather than opaque predictions. By combining this rigor in ML design with thermodynamic embedding, the proposed methodology represents a substantively advanced framework that moves beyond prior empirical, DoE-based, or unconstrained ML efforts.

In summary, compared with existing predictive tools, ML scales to complex, nonlinear, high-dimensional problem settings and yields state-of-the-art predictive accuracy, while TDM provides physical fidelity, interpretability, and extrapolation safeguards. Their integration directly addresses the historical shortcomings of empirical/DoE methods (limited nonlinearity and transferability) and CFD-only workflows (computational cost and calibration sensitivity), delivering a cohesive, reliable, and scalable basis for rational AFB design and engine calibration (Khan et al., 2020; Chia et al., 2020; Roy et al., 2021; Zhang et al., 2022; Sahin et al., 2022; Moran et al., 2015; Heywood, 2018; Turns, 2013; Onwusa et al., 2025).

A clear conceptual gap therefore persists in the literature. Current optimization studies tend to adopt either empirical or thermodynamic approaches that lack flexibility and scalability, or ML-driven strategies that sacrifice physical interpretability and experimental grounding. Although DEE has been repeatedly used in ML-based predictive studies, these works remain fuel-specific, fragmented and insufficiently generalizable, underscoring the absence of a cohesive

framework that can systematically unify DEE and other oxygenated fuels within a broader optimization strategy. This disjunction significantly restricts the transferability of results across fuel types, operating conditions and real-world driving scenarios, leaving the field without a cohesive and scalable optimization tool for AFBs. The novelty of the present study lies in its development of an integrated hybrid optimization framework that unifies ML, TDM and EVMs for alternative fuel blend optimization (FBO). Unlike previous fragmented approaches, the proposed methodology offers four distinct contributions. First, it bridges predictive analytics with physical laws by embedding thermodynamic constraints into ML algorithms, ensuring mechanistic interpretability. Second, it provides experimental anchoring through laboratory validation, thereby reinforcing practical relevance. Third, it enhances scalability by leveraging a curated dataset of diverse AFB compositions and operating regimes, enabling predictions that extend beyond single-fuel or engine-specific contexts. Finally, it introduces a structured triadic synergy (ML–TDM–EVMs) for rational fuel design, offering a unified pathway to optimize next-generation fuels such as advanced biofuels, synthetic hydrocarbons and alcohol-ether blends. Collectively, these contributions move beyond conventional empirical, theoretical, or purely data-driven studies, establishing a robust and adaptable optimization strategy

In conclusion, this study proposes a novel computational-experimental paradigm for fuel blend optimization in ICEs, integrating ML, thermodynamic principles and experimental testing into a single cohesive framework. By combining predictive precision with physical grounding and empirical validation, the methodology addresses a long-standing gap in the field the absence of interpretable and scalable optimization models. The insights from this research are expected to provide a versatile tool for scientists, engineers, and policymakers, ultimately supporting the global transition toward carbon-neutral, energy-secure, and high-efficiency transportation systems.

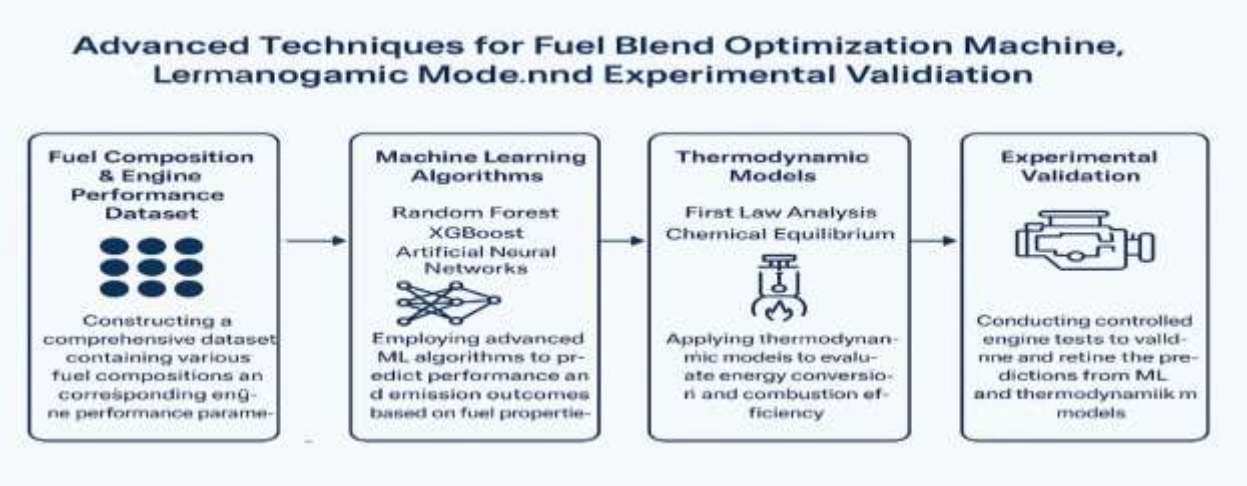

Figure 1: Visual abstract summary of advanced techniques for fuel blend optimization using ML, thermodynamic modeling and experimental validation methods

2.0. Materials and Methods

2.1. Materials

2.1.2. Fuel Sample Preparation and Characterization

A comprehensive set of fuel blends was prepared by combining base diesel with a selection of alternative fuel components, including fatty acid methyl esters (FAME), bioethanol, n-butanol, diethyl ether, and synthetic hydrocarbons. These alternatives were chosen based on their distinct physicochemical characteristics that influence combustion behaviour, such as oxygen content, cetane number, volatility, and heating value.

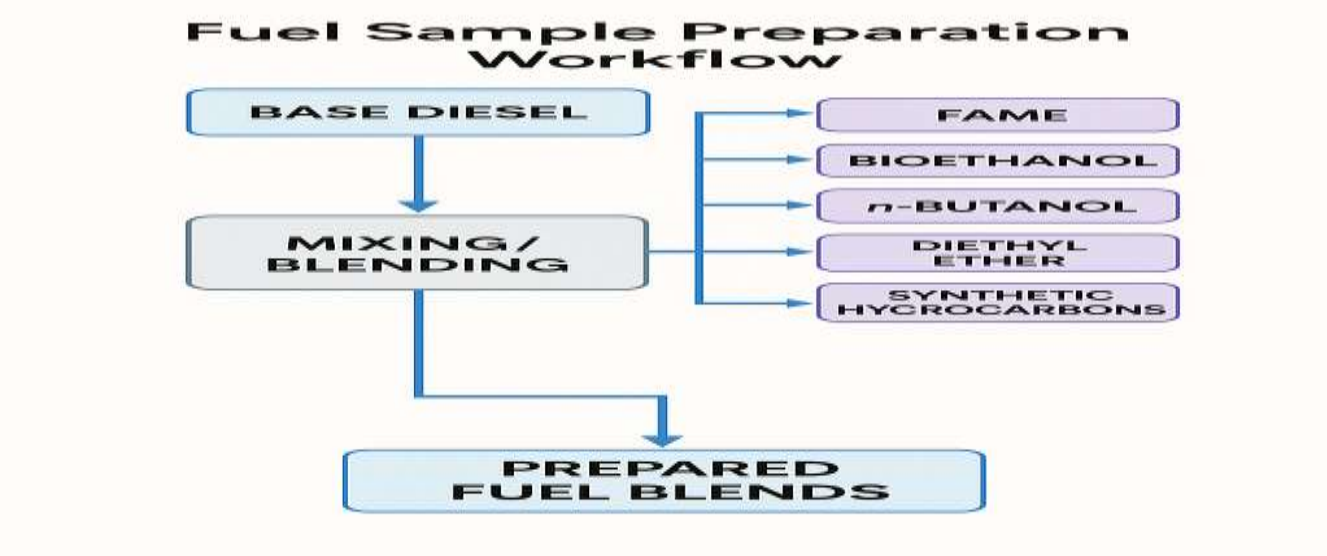

Figure 2: Fuel sample preparation workflow

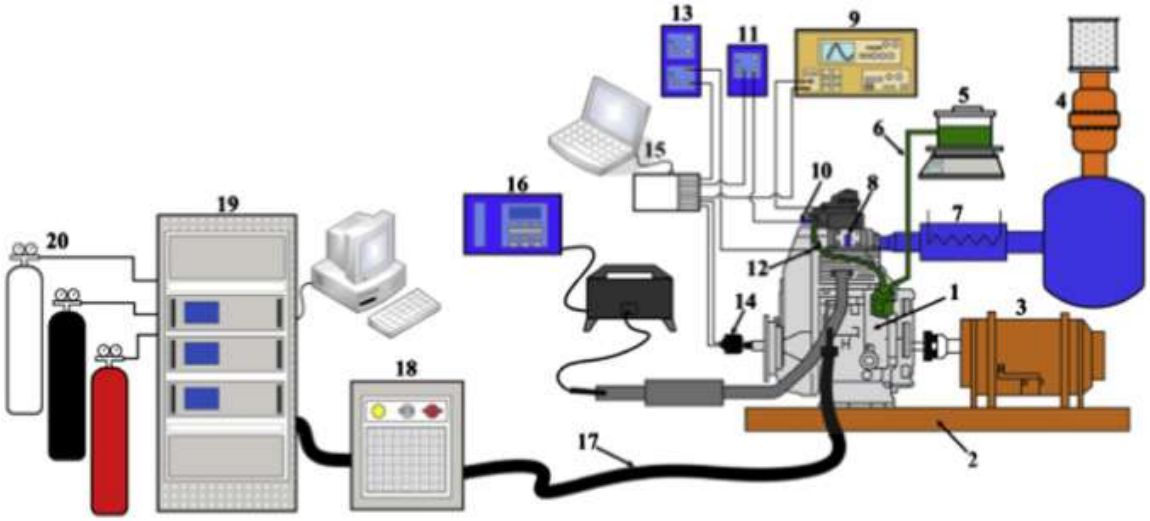

Figure 3: Schematic diagram of blending fuel mode system

This Figure 3 show a schematic layout of an engine test bed setup used for performance and emission analysis. Now is the accurate labeling of each numbered component in the diagram:

1. Internal Combustion (IC) Engine
2. Engine Test Bed / Base Frame
3. Dynamometer (Loading Device)
4. Dynamometer Cooling Water Circuit
5. Air Intake Manifold with Air Box
6. Air Flow Measuring Device (Air Drum / Orifice Meter / Airbox with Manometer)
7. Exhaust Silencer
8. Fuel Tank
9. Data Acquisition System (DAQ) / Signal Conditioner
10. Fuel Flow Measuring Unit
11. Cooling Water Heat Exchanger
12. Exhaust Gas Sampling Port
13. Exhaust Gas Analyzer (CO, HC, NOx, CO_2, O_2 measurements)
14. Crank Angle Encoder / TDC Encoder

15. Interfacing Module (Signal Processing)
16. Computer with Engine Performance Software
17. Wiring Harness & Sensor Cables
18. Dynamometer Control Panel (Load Controller)
19. Emission Analyzers Rack / Control Console
20. Calibration Gas Cylinders (CO_2, CO, NOx, O_2 mixtures)

This setup enables measurement of engine performance (torque, power, fuel consumption, efficiency) and emission characteristics.

2.1.3. Blend Composition

Both binary and ternary fuel blends were formulated, with alternative components mixed in volumetric ratios ranging from 10% to 50%. To ensure operational compatibility, cetane improvers (2-ethylhexyl nitrate) and anti-corrosion additives were incorporated in specific blends. A total of 12 unique fuel categories were created for further testing. Figure 4 show the diagram illustrates the production and utilization process of ternary biodiesel blends derived from *Ambadi seeds and waste cooking oil,* followed by performance evaluation in an engine.

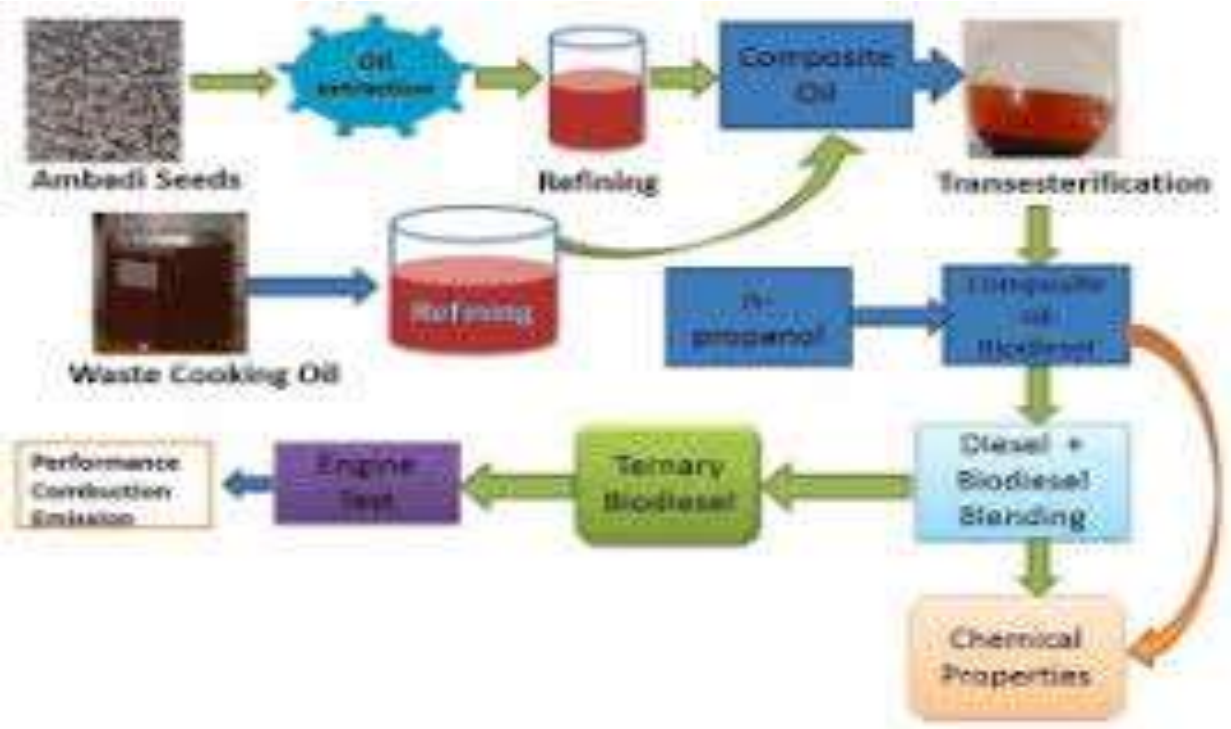

Figure 4: Blend composition

This diagram in figure 4 illustrates the production and utilization process of ternary biodiesel blends derived from Ambadi seeds *and* waste cooking oil, followed by performance evaluation in an engine. Here's a brief explanation step by step:

1. **Feedstocks**:
    a. **Ambadi Seeds** undergo oil extraction to produce crude oil.
    b. **Waste Cooking Oil** is also collected and subjected to refining to remove impurities.
2. **Refining & Composite Oil Formation**: Both extracted Ambadi seed oil and refined waste cooking oil are combined to form a composite oil**.**
3. **Transesterification Process**:
    a. The composite oil reacts with n-propanol (alcohol) in the presence of a catalyst.
    b. This chemical process, known as transesterification**,** converts the composite oil into biodiesel (methyl/propyl esters) and glycerol by-product.
4. **Blending Stage**:
    a. The resulting biodiesel is blended with diesel fuel to obtain various proportions (biodiesel–diesel blends).

b. **Ternary biodiesel** refers to a mixture prepared using three components: diesel, Ambadi biodiesel, and waste cooking oil biodiesel.

5. **Property Analysis**: The chemical properties (density, viscosity, calorific value, etc.) of the produced blends are evaluated.
6. **Engine Testing**:
   a. The ternary biodiesel blends are tested in an engine test setup**.**
   b. The study evaluates performance (brake power, efficiency, fuel consumption)**,** combustion characteristics and emission parameters (CO, HC, NOx, smoke opacity, etc.)**.**

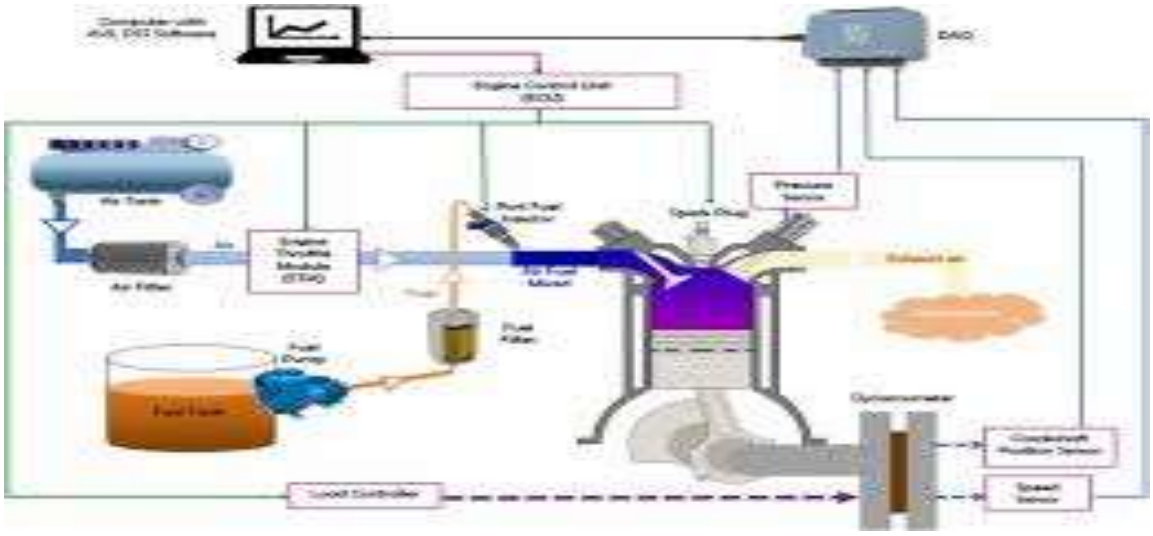

Figure 5: Ternary fuel blends

This diagram in figure 5 represents a computer-controlled engine test setup used to evaluate the performance, combustion, and emission characteristics of an internal combustion engine under controlled conditions. At this juncture is a brief explanation of the process flow:

1. **Air Supply System**
   i. **Air Tank & Air Filter**: Provide clean and filtered air to the engine.
   ii. Air passes through an Engine Throttle Module (ETM) to regulate the air intake.
2. **Fuel Supply System**
   i. **Fuel Tank → Fuel Pump → Fuel Filter**: Supplies clean pressurized fuel to the Fuel Injector.
   ii. The injector sprays the required quantity of fuel into the intake manifold/cylinder.
3. **Combustion Chamber**
   a. Inside the cylinder, air and fuel mix and combustion occurs, initiated by the spark plug (for SI engines).
   b. A **pressure sensor** monitors in-cylinder pressure for combustion analysis.
   c. Exhaust gases exit through the exhaust outlet for emissions measurement.
4. **Engine Control & Measurement**
   i. **Engine Control Unit (ECU)**: Governs injection timing, spark timing, and throttle position.
   ii. **Crankshaft Position Sensor**: Tracks piston position and engine speed.
   iii. **Dynamometer**: Provides mechanical load on the engine to simulate real driving conditions.
   iv. **Load Controller**: Adjusts the applied load for performance evaluation.
5. **Data Acquisition System (DAQ)**
   i. All sensor data (fuel flow, air flow, in-cylinder pressure, speed, torque, exhaust conditions) are sent to the DAQ.
   ii. A computer with AVL/DEW software records and analyzes engine parameters such as brake power, brake thermal efficiency, specific fuel consumption, and combustion/emission data

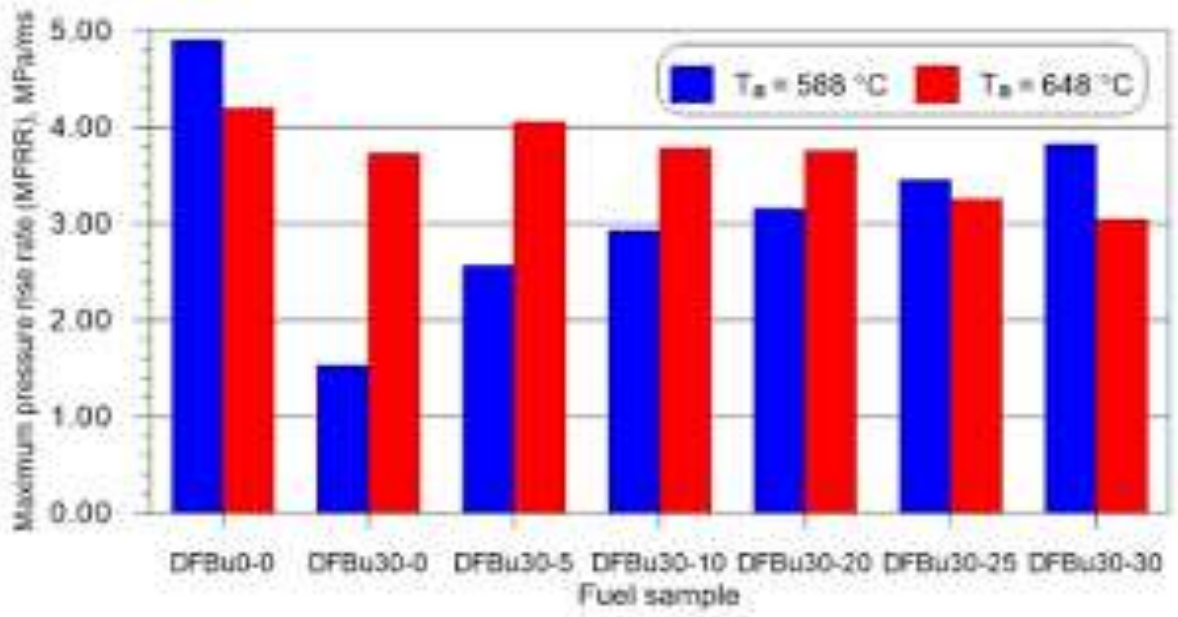

Figure 6a: Bar chart of Cetane improvers (2-ethylhexyl nitrate)

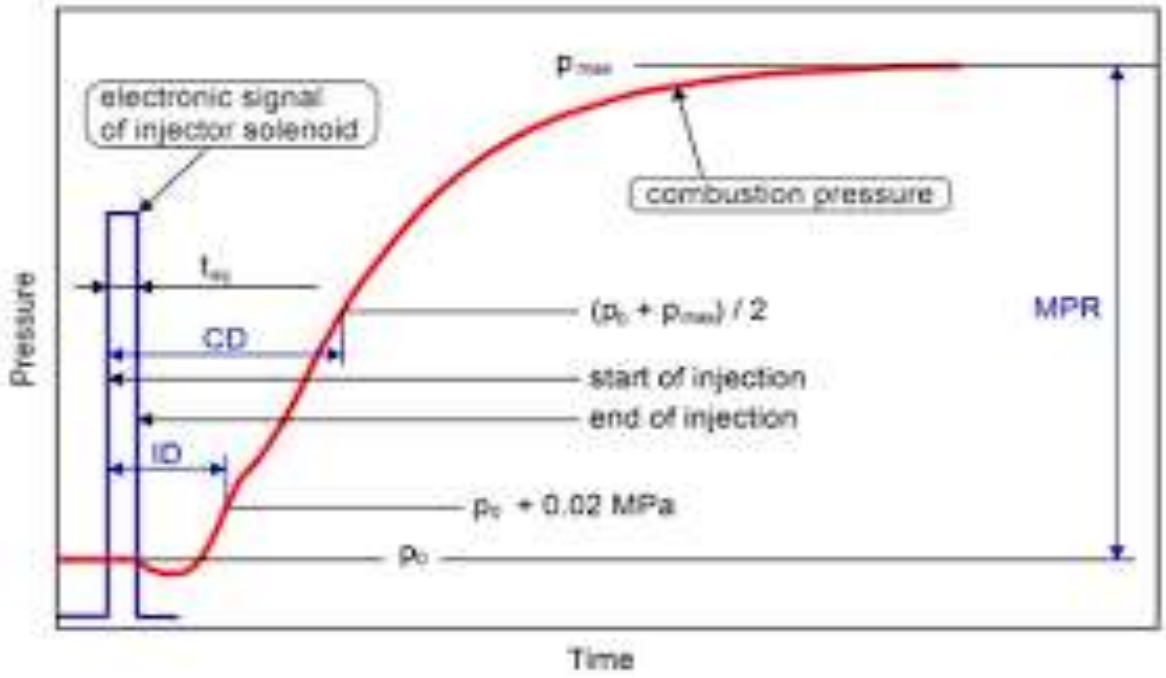

Figure 6b: Graphical of Cetane improvers (2-ethylhexyl nitrate)

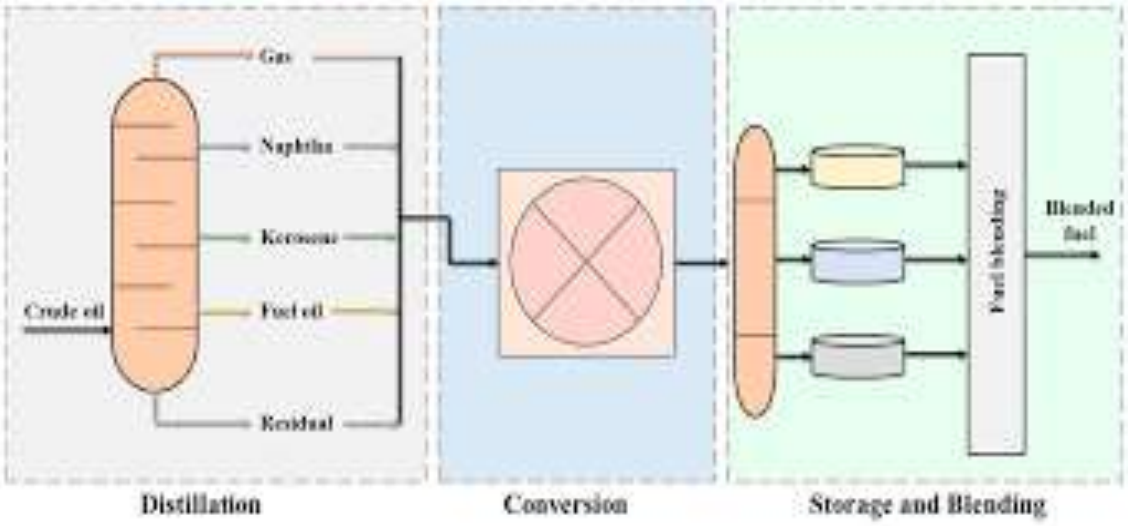

Figure 7: Fuel blending process

This diagram in figure 7 illustrates the fuel production process from crude oil refining to the final blended fuel. It is divided into three main stages:

1. **Distillation**: Crude oil is heated and separated into different fractions based on boiling points. The main products are:
   i. **Gas** (lightest fraction)
   ii. **Naphtha** (used in gasoline production)
   iii. **Kerosene** (jet fuel and heating)
   iv. **Fuel oil** (diesel and heavy oils)
   v. **Residue** (asphalt, heavy oils)
2. **Conversion**
   i. Heavier fractions are chemically processed (e.g., cracking, reforming) into lighter and more valuable products such as gasoline and diesel.
   ii. This step increases fuel yield and improves quality.
3. **Storage and Blending**
   a. Refined products are stored in tanks.
   b. Blending is carried out to meet specific fuel standards (e.g., octane rating, sulfur content).
   c. The result is the blended fuel (gasoline, diesel, etc.) supplied to consumers.

Table 1: Fuel property measurements and corresponding standards

S/N	Fuel Property	Test Method (Standard)	Description/Measurement Approach
1	Cetane Number	ASTM D613	Indicates combustion quality; higher values signify better ignition quality.
2	Kinematic Viscosity	ASTM D445	Measures the fuel's resistance to flow under gravity at a given temperature.
3	Density	ASTM D4052	Determines mass per unit volume, affecting atomization and energy content.
4	Lower Heating Value (LHV)	ASTM D240	Quantifies the energy released during complete combustion of the fuel.
	Oxygen Content	CHN Elemental Analysis	Measures the percentage of oxygen in the fuel, influencing combustion efficiency and emissions.

These parameters formed the foundational input features for the ML model

2.1.4. Data Collection and Dataset Construction

A structured dataset was compiled through systematic experimental testing, serving as a training and validation base for machine learning models (MLMs).

2.1.5. Engine Test Bench Setup

A single-cylinder, four-stroke, direct-injection diesel engine was employed to evaluate the performance and emission characteristics of each fuel blend. The engine was operated under controlled conditions, varying across three different load levels (25%, 50%, and 100%) and engine speeds (1500 rpm, 2000 rpm, and 2500 rpm).

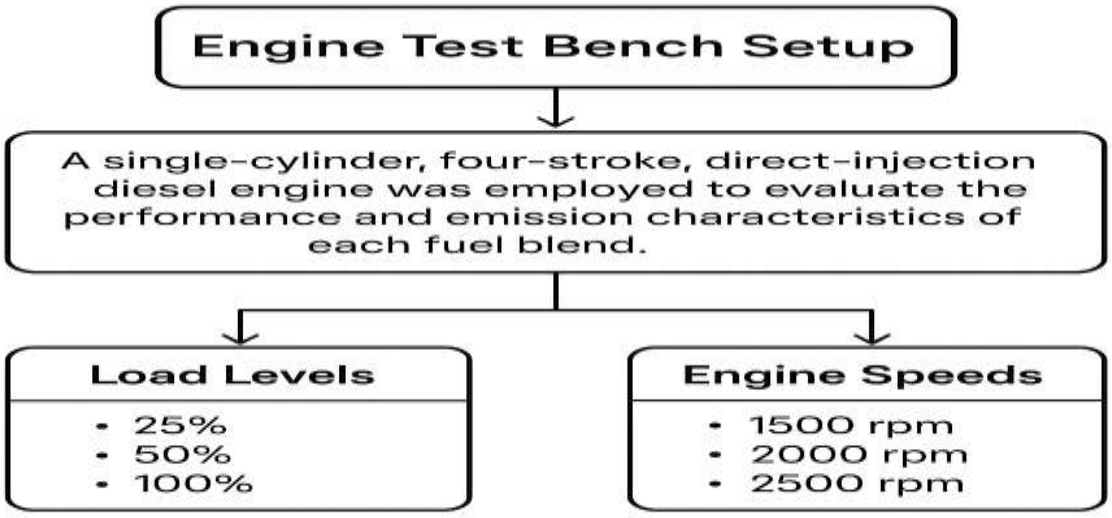

Figure 8: Flowchart showing the engine test bench setup

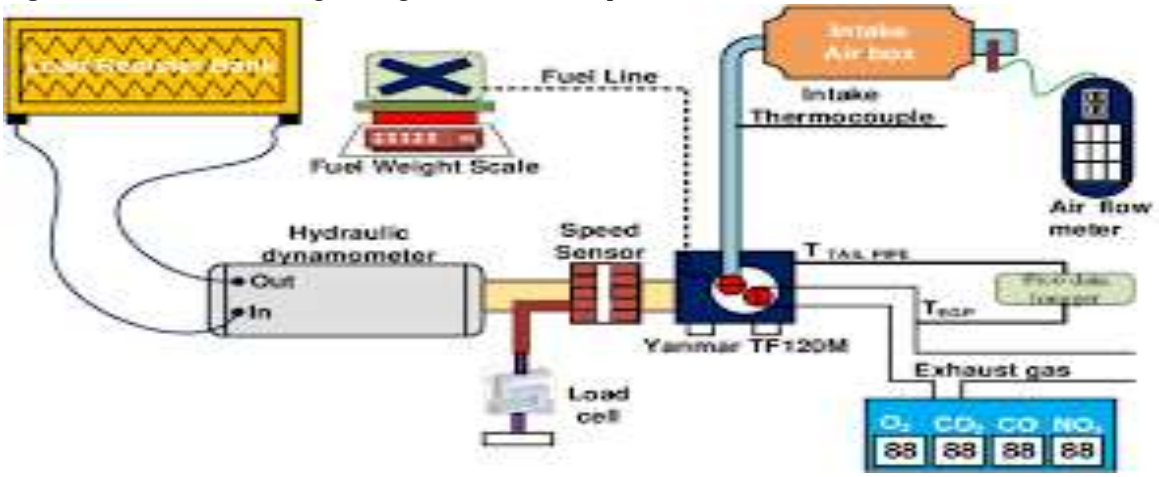

Figure 9a: A single-cylinder diesel engine

This in figure 9a diagram represents an engine test setup for measuring performance and emissions of a diesel engine (Yanmar TF120M). Here's a brief explanation of the main components and flow:

1. **Intake System**
   i. **Air box & Air flow meter**: Measure and regulate the air entering the engine.
   ii. **Intake thermocouple**: Records intake air temperature.
2. **Fuel Supply System:** Fuel weight scale: Monitors fuel consumption by measuring the mass of fuel supplied through the fuel line.
3. **Engine (Yanmar TF120M):** Connected to the test rig, where combustion occurs using measured air and fuel inputs.
4. **Exhaust Measurement**
   i. **Thermocouples** (T_pipe, T_exh)**:** Measure exhaust gas temperatures.

ii. **Exhaust gas analyzer**: Detects pollutant levels (O_2, CO, CO_2, NO, NO_2).

5. **Performance Measurement**
   i. **Hydraulic dynamometer**: Provides engine load by absorbing power.
   ii. **Load cell**: Measures applied load/torque.
   iii. **Speed sensor**: Records engine speed.
6. **Cooling & Load Absorption:** The dynamometer uses circulating water to absorb engine power and maintain controlled operating conditions.

Purpose: This setup enables researchers to evaluate engine performance (power, torque, fuel consumption) and emission characteristics (O_2, CO, CO_2, NOx) under controlled conditions.

This diagram in figure 9b shows a diesel engine test rig commonly used in laboratories for performance and emission testing.



Figure 9b: four-stroke diesel engine bench test

2.1.6. **Main Components and Functions:**

1. **Diesel Engine (red unit on the right)**
   i. The prime mover under test, usually a single-cylinder diesel engine.
   ii. Provides power output that is evaluated for performance (torque, speed, efficiency).
2. **Dynamometer (blue unit coupled to the engine)**
   i. Connected directly to the engine shaft.
   ii. Applies a controlled load to the engine.
   iii. Measures torque and power output.
3. **Control and Measurement Panel (white vertical unit with instruments)**
   i. Houses sensors and meters for measuring key parameters such as speed, torque, fuel consumption, and temperature.
   ii. Displays real-time data for monitoring engine performance.
4. **Cooling and Fuel System (tank, piping, gauges)**
   i. Circulating water system to cool the engine and dynamometer.
   ii. Fuel supply system for controlled delivery and measurement of fuel consumption.
5. **Base Frame (blue support structure):** Rigid platform where the engine and dynamometer are mounted, ensuring proper alignment and stability.

Purpose: This setup is used for:

i. **Performance testing**: Brake power, brake thermal efficiency, brake specific fuel consumption.
ii. **Emission testing** (with analyzers attached): CO, CO_2, NOx, etc.
iii. **Educational and research applications**: To study combustion, alternative fuels, and engine behaviour under various loads.

2.1.7. Instrumentation and Measurements

i. **In-cylinder pressure** was recorded using a high-fidelity piezoelectric transducer, synchronized with a crank angle encoder.
ii. **Exhaust gas composition** (NO_x, CO, HC) was analyzed using an AVL DIGAS 444 gas analyzer.
iii. **Smoke opacity** was measured using a Hartridge smoke meter.
iv. A **LabVIEW-based data acquisition system** enabled real-time logging and post-processing.

2.1.8. Dataset Structure

The experimental dataset comprised just over 500 cases, each recording both fuel blend characteristics and engine performance responses. Although this dataset size is modest for training artificial neural networks (ANNs), careful pre-processing and validation protocols were applied to mitigate over fitting and enhance model robustness. Each record included:

a. **Input variables**: Fuel blend composition and physicochemical properties (cetane number, viscosity, density, oxygen content, volatility, and lower heating value).
b. **Output responses**: Brake Thermal Efficiency (BTE), Brake Specific Fuel Consumption (BSFC), NO_x, CO, HC, and smoke opacity.

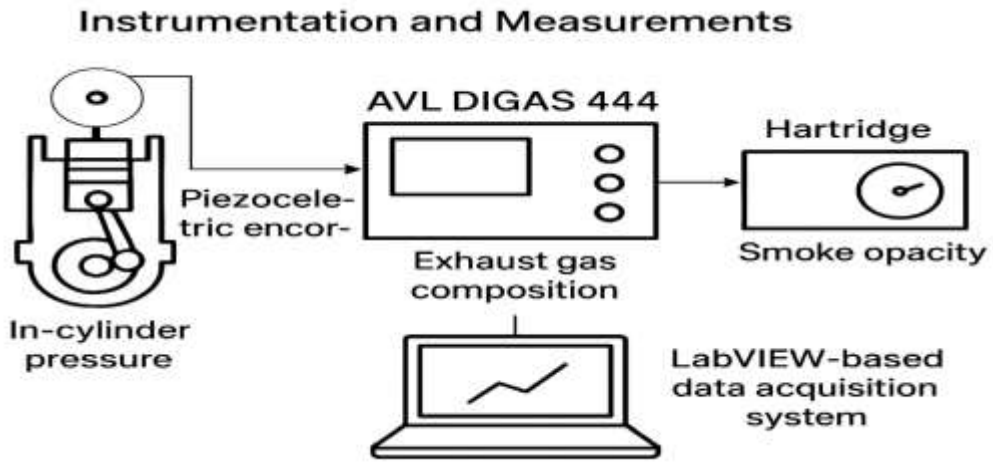

Figure 10: Instrumentation and Measurements

2.1.9. Machine Learning Modeling (MLM): MLMs were developed to predict engine responses based on input fuel properties, capturing complex nonlinear interactions that are difficult to model analytically.

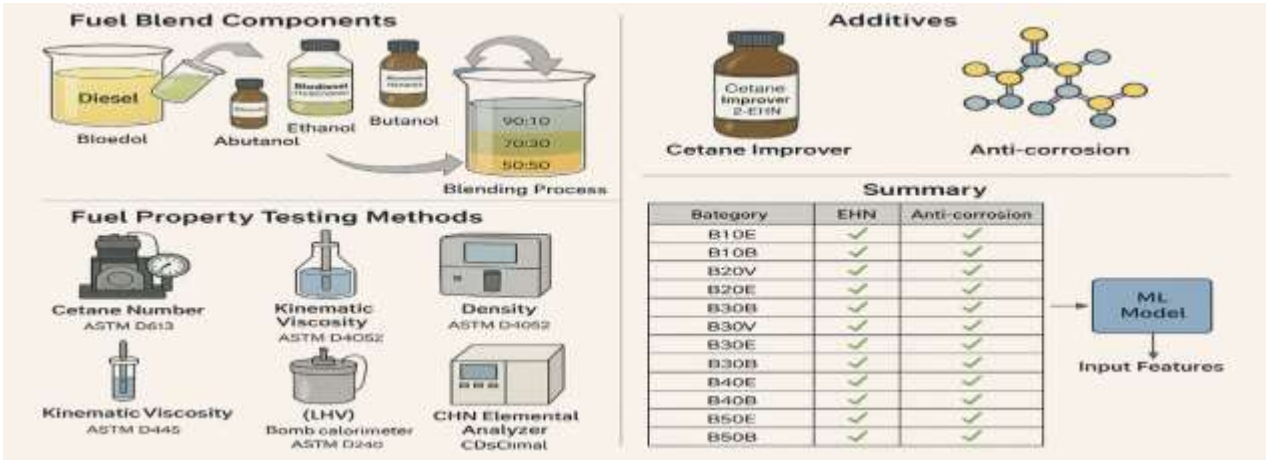

Figure 11: This illustration presents a comprehensive overview of a fuel blend evaluation process used for developing and analyzing biofuel-based alternative fuel blends enhanced with additives, tested using standard fuel property testing methods and evaluated through ML models.

2.1.10 Preprocessing and Feature Engineering

To ensure uniformity and comparability, all input variables were normalized using z-score standardization. Pearson correlation analysis was conducted to identify strongly correlated variables and PCA was applied to reduce dimensionality, retaining the most significant orthogonal components that dominantly influenced emissions and efficiency outcomes.

2.1.11 Model Training and Validation Protocols

Three ML models RF, XGBoost and ANNs were implemented. Compared with earlier studies that applied these models without justification, the present study adopted a structured workflow to ensure fairness, interpretability and generalizability:

a. **Dataset splitting**: The dataset was divided into training (70%), validation (15%), and testing (15%) subsets, with randomization controlled by fixed seeds to ensure reproducibility.

b. **Cross-validation**: A 5-fold cross-validation strategy was applied within the training set to prevent over fitting and ensure that models generalized across diverse blend–load combinations.
c. **Over fitting mitigation**: Regularization (L2 for RF/XGBoost, dropout layers for ANNs), early stopping criteria, and hyperparameter optimization (grid/random search) were employed to control variance and avoid memorization of small-sample artifacts.
d. **Hyperparameter tuning**: RF parameters (number of trees, max depth), XGBoost boosting iterations, and ANN architecture (hidden layers, learning rate, batch size) were systematically tuned using validation error as the selection criterion.

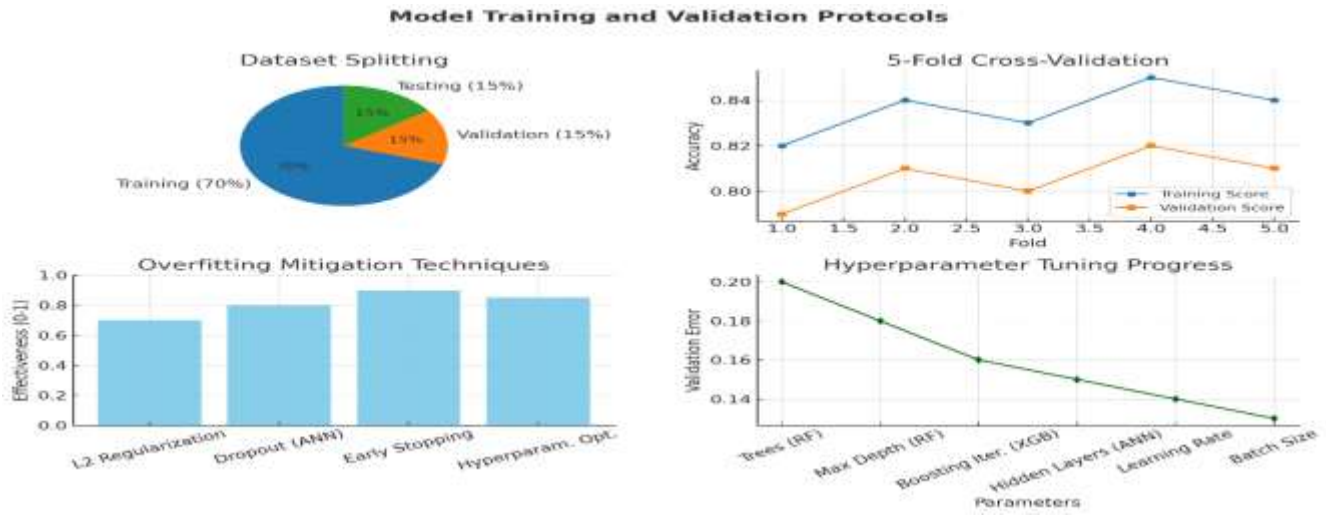

Figure 12: Model training and validation protocols

2.1.12 Addressing Data Scarcity in ANN Training

Given that 500 experimental cases is relatively small for ANN training**,** strategies were adopted to improve robustness:

a. Compact network architectures (fewer hidden layers/neurons) were used to minimize over-parameterization.
b. Weight regularization and dropout were applied to stabilize training.
c. Bootstrapping and data augmentation (interpolated blends within thermodynamically consistent limits) were explored to increase the effective learning base without compromising physical fidelity.

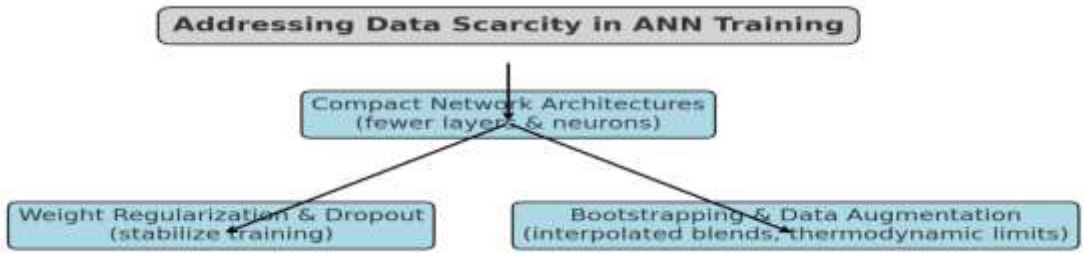

Figure 13: Addressing Data Scarcity in ANN Training*:*

The central title flows downward into three strategy boxes:

1. **Compact Network Architectures** → fewer layers/neurons to avoid over-parameterization.
2. **Weight Regularization and Dropout** → stabilizes training.
3. **Bootstrapping and Data Augmentation** → generates interpolated blends within thermodynamic limits.

2.1.13 Model Evaluation and Interpretability

Model performance was evaluated using RMSE, MAE and R^2 across both validation and independent test sets. Beyond predictive accuracy, interpretability tools including permutation importance, partial dependence plots and SHAP value analysis were applied to reveal the relative contributions of fuel properties and operating conditions to performance and emissions. This ensured the models were not only predictive but also physically meaningful

2.1.14. Machine Learning Innovation

1. **Hybrid Learning Framework with Physics-Informed Constraints**

a. Instead of relying purely on black-box ML, the models incorporate thermodynamic priors -conservation of energy, mass balance, entropy bounds into loss functions.
b. This Physics-Informed Machine Learning (PIML) approach reduces over fitting on small datasets, enhances interpretability, and ensures that predictions remain physically valid.
c. Novelty: Moves beyond empirical ML by embedding first-principles knowledge, ensuring generalizability across blend–load–speed domains.

2.1.15. Thermodynamic Modeling (TDM)

TDMs were constructed in parallel with ML to ensure that predictions adhered to physical laws and combustion theory

2.1.16. First Law of Thermodynamics: The energy balance of the combustion process was modeled using in-cylinder pressure data to calculate: Heat release rate, indicated power and indicated thermal efficiency, A Wiebe function was fitted to model the combustion phase over crank angle degrees.

Equations $-\frac{dQcomb}{\delta\theta} = \frac{\Upsilon}{\Upsilon-1} P \frac{dv}{d\theta} + \frac{1}{\Upsilon-1} V \frac{dP}{\delta\theta}$ (1)

Where;

$\frac{dQ}{d\theta}_{comb}$: Rate of heat release w.r.t crank angle

Υ: Specific heat ratio

P: Instantaneous cylinder pressure

V: Instantaneous cylinder volume

θ: Crank angle

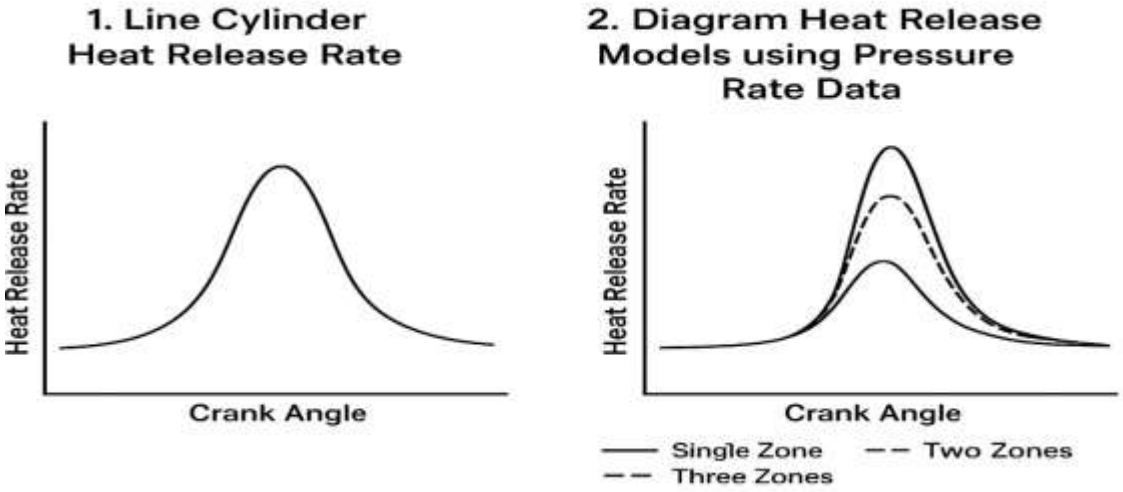

Figure 14; Line cylinder heat release rate and 2. Diagram heat release models using pressure rate data

2.1.17. **Combustion Modelling using Wiebe Function**

To represent to mass fraction burned (MFM) during combustion

Wiebe function equation: $xb(\theta) = I - \exp\left[-a\left(\frac{\theta-\theta_0}{\Delta\theta}\right)^m\right]$ (2)

Where:

xb(θ): mass function burned

Δθ: combustion duration

θ: start of combustion

a, m; empirical shape presenters

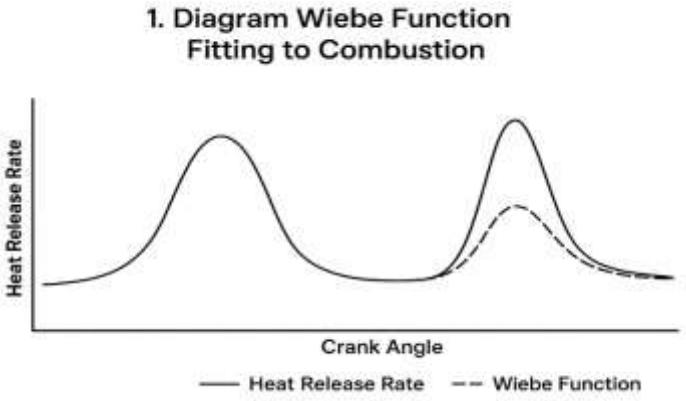

Figure 15: 1. Diagram of Combustion Modelling using Wiebe Function

2.1.18. **Indicated Power and Efficiency**

Using heat release data and cylinder; pressure volume integration

Indicated Power: $IP = \frac{1}{2\pi} \int_{cyde}^{Pdv}$ (3)

Indicated Thermal Efficiency: $\eta thi = \frac{IP}{\dot{m}f - LHV}$ (4)

Where;

ηthi: fuel mass flow rate

LHV: Lower heating value of the fuel

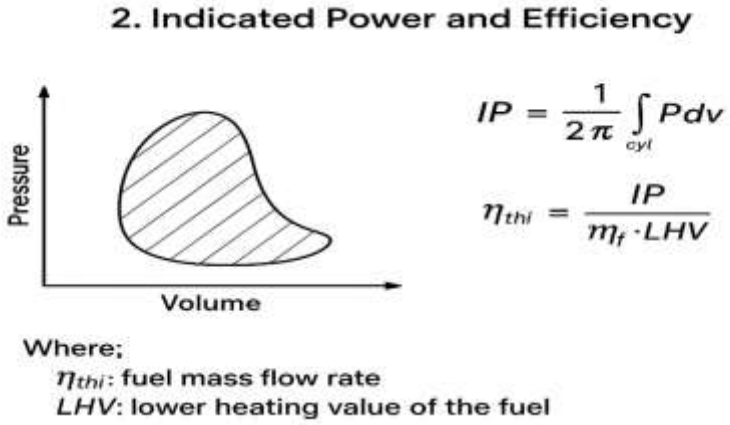

Figure 16: Indicated Power and Efficiency

2.1.19. Second Law and Exergy Analysis (EA): The Second Law of Thermodynamics was applied to compute entropy generation and irreversibilities during the combustion process. Exergy efficiency was also calculated to assess the quality of energy conversion. Second Law and Energy Analysis (EA) Entropy Generation (S_{gen})

$S_{gen} = \Delta S_{system} + \Delta S_{surrounding}$ (5)

High entropy generation = high irreversibility

Energy efficiency

$\eta ex = \frac{useful\ energy\ output}{exergy\ input} = \frac{W\ useful}{E\ fuel}$ (6)

Where;

E_{fuel}: ṁf – ex_{fuel}

W_{useful}: Work extracted

Ex_{fuel}: Specific energy fuel

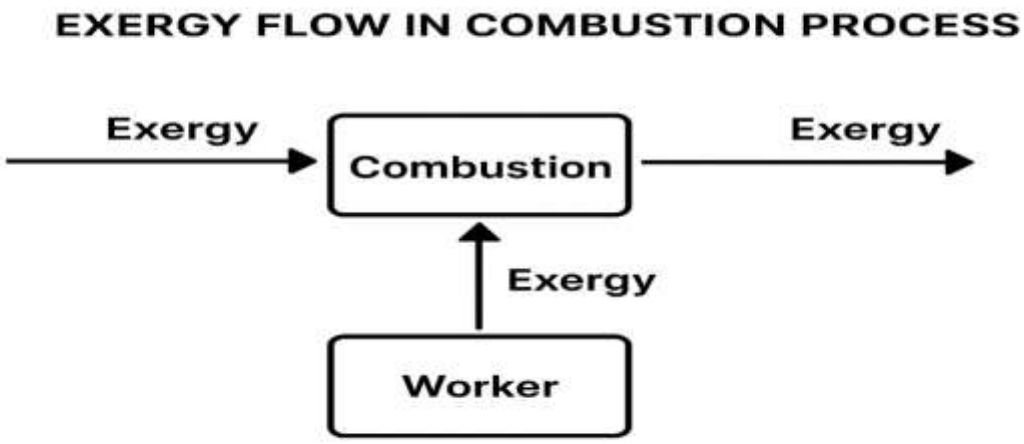

Figure 17; Exergy flow in combustion process

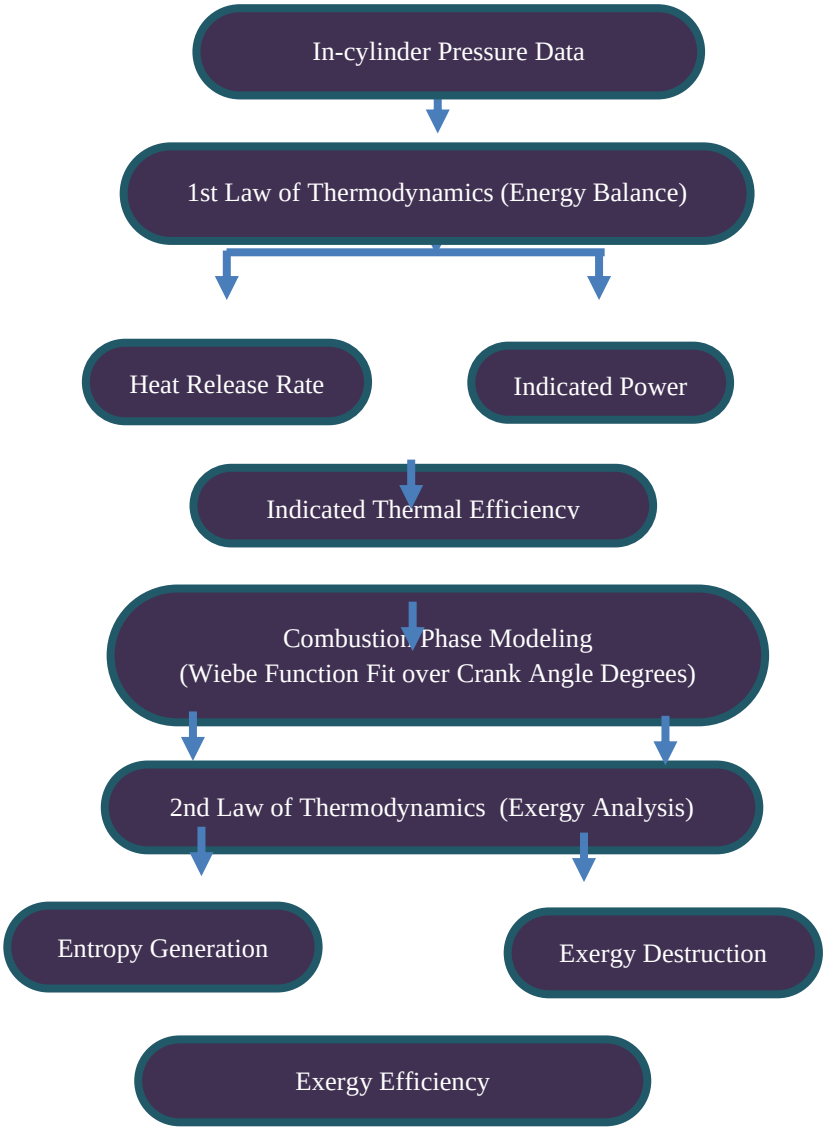

Figure 18: A flow chart of TDM

2.1.17. Chemical Equilibrium Analysis: Chemical equilibrium modeling was performed using the NASA CEA code to estimate theoretical emissions, flame temperatures, and equilibrium species concentrations. This analysis established the physical bounds for NO_x, CO, and unburned hydrocarbon predictions, thereby improving the interpretability of ML outputs.

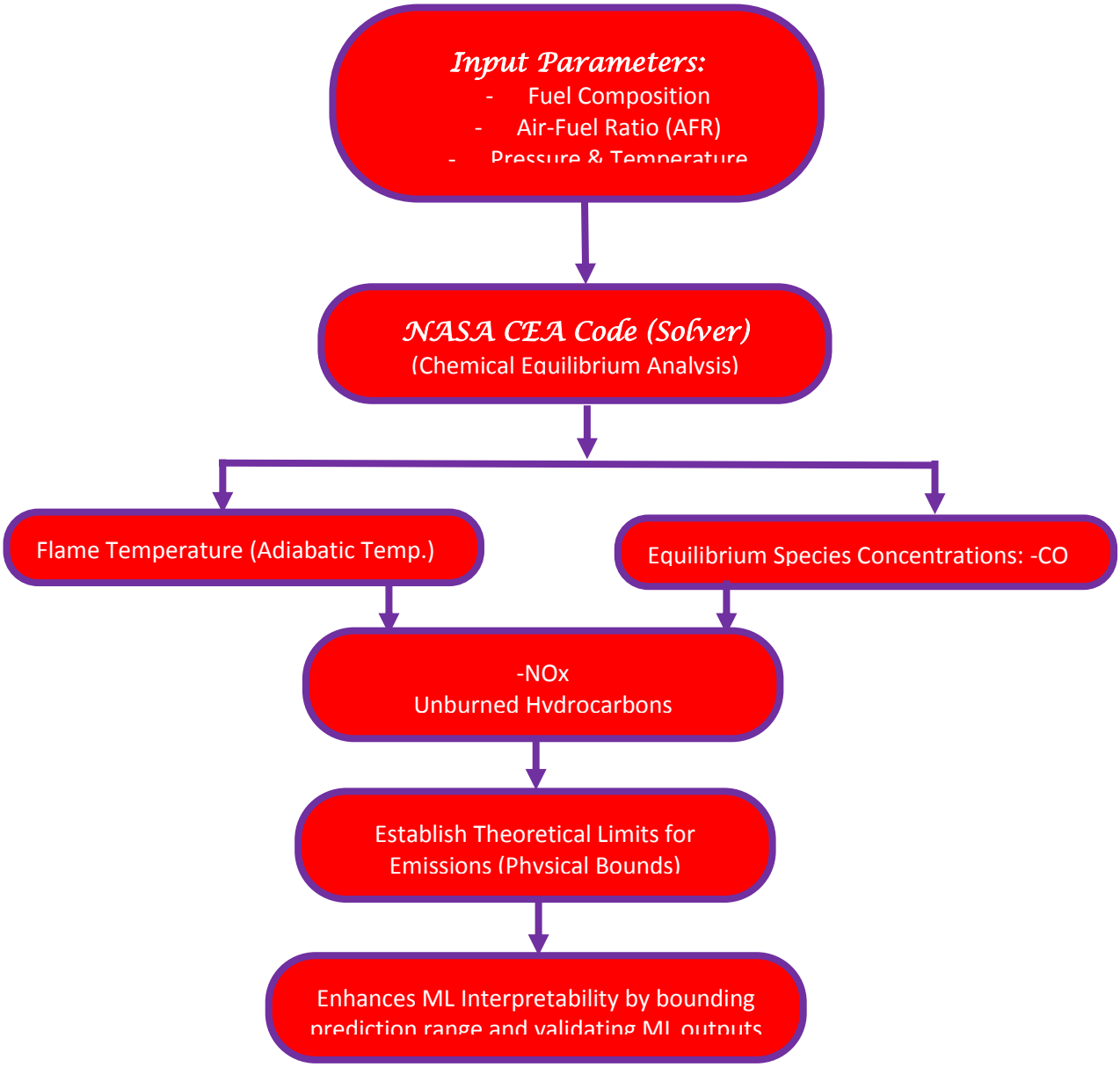

Figure18: Chemical equilibrium analysis

2.1.20. Model Integration and Multi-objective Optimization

2.1.21. Hybrid Framework Implementation: The outputs from MLMs were cross-validated against thermodynamic predictions. Any deviations violating energy conservation or stoichiometric constraints were flagged and corrected. This ensured physical reliability and improved model robustness.

2.1.22. Optimization Strategy: A multi-objective optimization approach was employed using the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to identify Pareto-optimal solutions that:

i. Maximized Brake Thermal Efficiency

ii. Minimized NO_x, CO, and HC emissions

The most promising blends from the Pareto front were shortlisted for empirical validation.

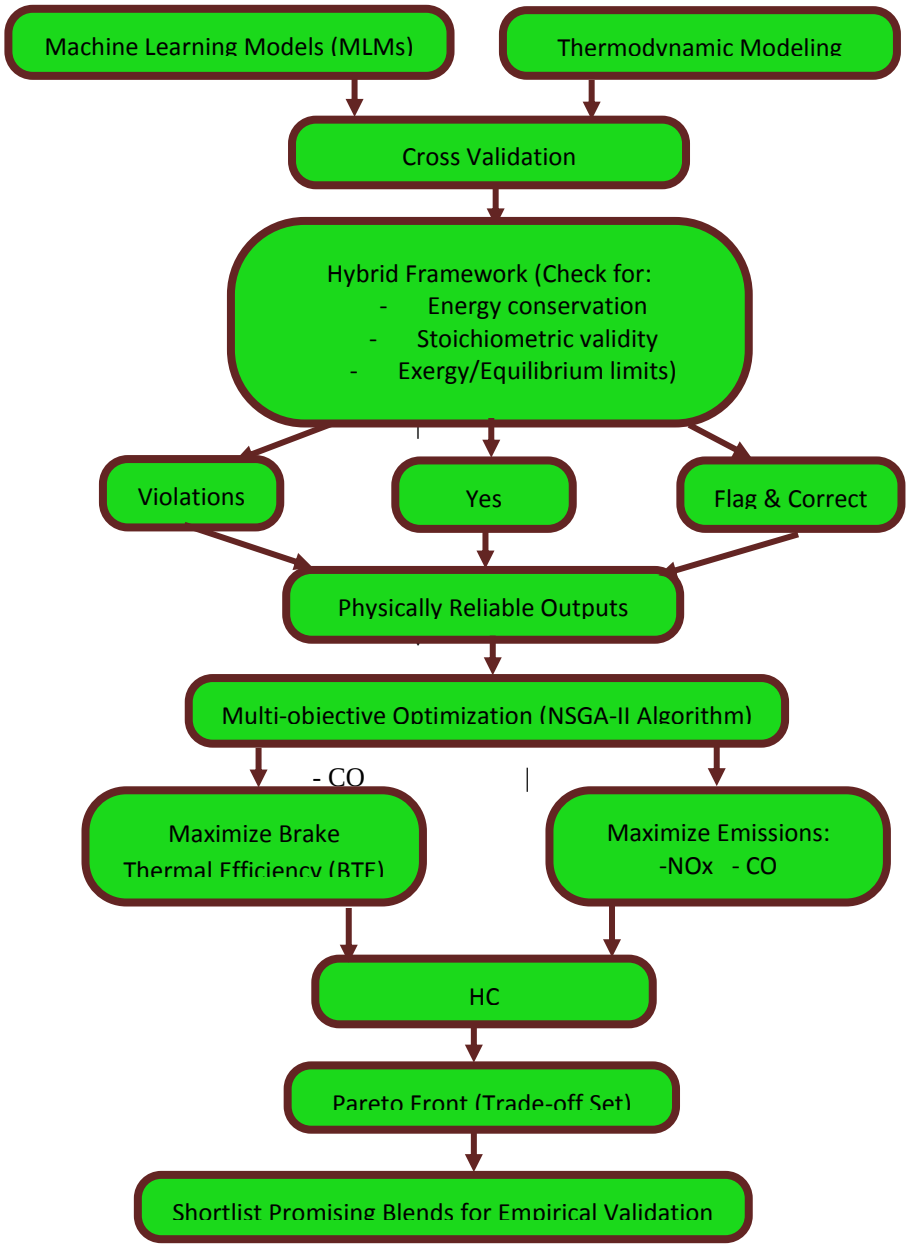

Figure19: MLMs and TDM

2.1.23. Thermodynamics Innovation

2.1.24. Thermodynamically Consistent Data Augmentation for Fuel Blends

a. Traditional interpolation of blend properties often violates thermodynamic limits. Here, augmentation is performed within Gibbs free energy and enthalpy consistency constraints, ensuring interpolated blends remain realistic.
b. This increases the effective dataset size without introducing unphysical points.
c. Novelty: Provides a physically rigorous augmentation strategy for data-scarce ANN training in combustion studies.

2.1.25. Experimental Validation

2.1.26. Independent Testing: A subset of 10 fuel blends, not used in model training, was experimentally tested under identical engine conditions. The measured engine responses were then compared with ML-thermodynamic predictions.

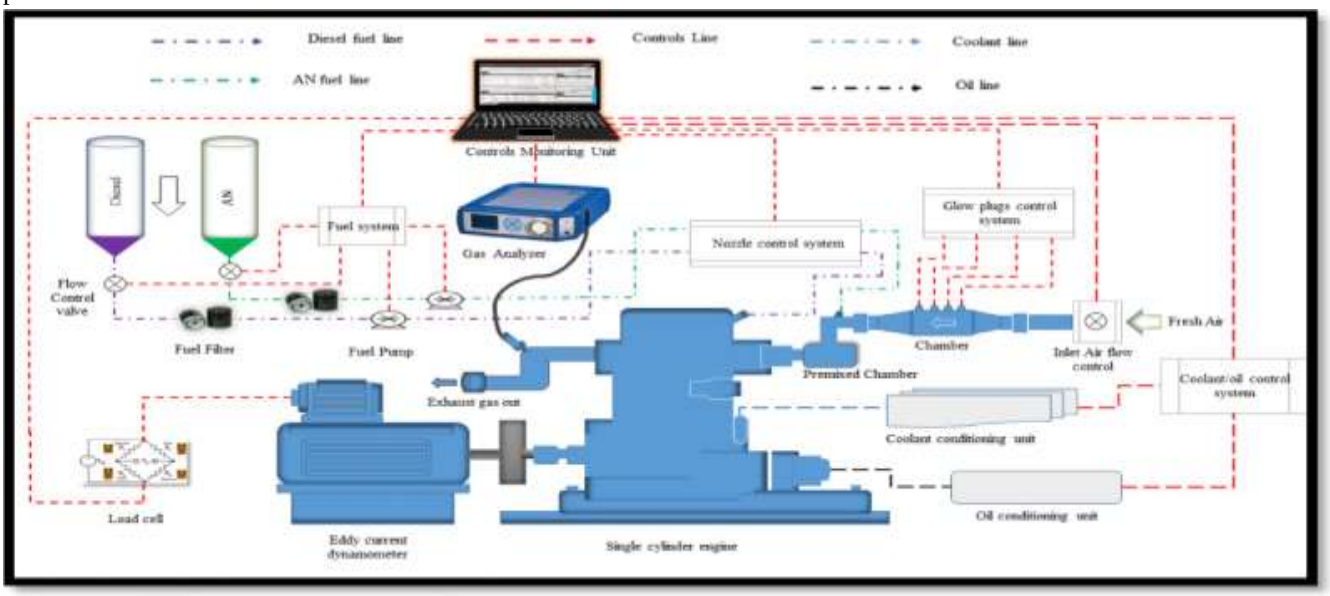

Figure 20: Experimental setup and validation of single-cylinder engine experimental test

This diagram in figure 20 represents a single-cylinder engine experimental test setup used to study fuel performance, combustion, and emissions. It integrates different subsystems for fuel supply, air intake, cooling, lubrication, control, and measurement. The key sections:

1. **Fuel Supply System**
   - i. Two fuel tanks are shown: Diesel and an alternative fuel (AN).
   - ii. Flow control valves and fuel filters regulate and clean the fuel before entering the fuel pump.
   - iii. Fuel is then delivered to the engine for combustion.
2. **Air Intake System**
   - i. Fresh air enters through the inlet air flow control and passes into the chamber.
   - ii. A premixed chamber ensures proper mixing of air and fuel before combustion.
   - iii. Glow plug and nozzle control systems assist in ignition and injection timing.
3. **Engine System**
   - i. The single-cylinder engine is the core of the setup.
   - ii. Connected to an eddy current dynamometer that measures engine output power and torque.
   - iii. Load cells are used to monitor engine load.
   - iv. Exhaust gases exit through the exhaust line and are analyzed by a gas analyzer for emissions.
4. **Cooling and Lubrication Systems**
   - i. **Coolant conditioning system** regulates engine temperature.
   - ii. **Oil conditioning unit** manages lubrication and oil cooling.
5. **Control and Monitoring**
   - i. A control and monitoring unit (computer-based system) manages and records all parameters.
   - ii. It links to various control systems (fuel, nozzle, glow plug, coolant, etc.) via control lines.

2.1.27. Experimental Setup Overview

This setup integrates hardware components, control systems, and analytical devices to evaluate the performance and emissions of a single-cylinder engine using different fuel blends. Main components and flow

- **a. Fuel Blends**:
  - i. Two separate fuel tanks: Diesel and AN (Alternative/Nitrogenated) Fuel.
  - ii. Each fuel is directed through dedicated fuel lines (purple for diesel, green for AN fuel).
  - iii. Flow control valves, fuel filters and a fuel pump regulate and purify the fuel flow to the engine.
- **b. Combustion System**:
  - i. Fuel is injected into the single-cylinder engine with the help of the nozzle control system.

ii. Glow plug control system ensures optimal combustion during cold starts or low-temperature conditions.
iii. Fresh air is supplied via the inlet air flow control and passes into the premixed chamber and main chamber.

c. Measurement and Control:
a. A Control and Monitoring Unit (CMU) (connected to a laptop) oversees real-time data logging, control of fuel systems, nozzle timing, and glow plug operations.
b. An eddy current dynamometer measures engine torque and speed.
c. A load cell may be used for force measurement.
d. Exhaust gases are analyzed using a gas analyzer to quantify emissions like NOx, CO, CO_2, HC, and particulates.

d. Thermal Management:
a. Coolant lines and an oil conditioning unit are included to maintain stable engine temperature and lubrication.
b. A coolant conditioning unit ensures that temperature effects do not skew performance or emission data.

2.1.28. Independent Testing

Objective: Validate the performance of the ML-thermodynamic prediction model.

Procedure:
i. A subset of 10 fuel blends, which were not used in the model training, were tested.
The engine was operated under identical controlled conditions using the setup above.
ii. Outputs (e.g., brake thermal efficiency, specific fuel consumption and exhaust emissions) were recorded.
iii. The experimental results were then compared to predictions made by the machine learning–thermodynamic model to evaluate model accuracy and generalization

2.1.29. Significance of Setup for Experimental Validation

i. Ensures reliable, repeatable, and controlled testing conditions.
ii. Allows independent verification of model outputs.
iii. Provides comprehensive data for performance analysis, covering fuel system dynamics, combustion behaviour, emissions, and thermal management.

2.1.30. Error and Sensitivity Analysis: Prediction errors were quantified using Mean Absolute Percentage Error (MAPE), with results showing average accuracies exceeding 90% across all performance and emission metrics. Global sensitivity analysis was conducted using the Sobol method to assess the impact of individual input variables on model output.

2.1.31. Experimental Innovation

2.1.32. **Integrated High-Fidelity Multi-Sensor Framework**

a. Real-time coupling of in-cylinder pressure transducers**,** exhaust gas analysers' and smoke opacity meters with a LabVIEW-driven adaptive acquisition system**.**
b. Unlike static measurements, this setup enables dynamic synchronization with crank-angle resolution, capturing transient combustion events.
c. Novelty: Produces a multi-dimensional dataset (pressure–emission–opacity correlations) that can feed directly into ML pipelines for joint optimization

2.1.32. Mathematical Derivatives and Calculations

Advancing techniques for fuel blend optimization using machine learning, thermodynamic modelling and experimental validation methods

i).**Thermodynamic Modelling:** Thermodynamic properties are essential for modelling combustion, efficiency and emissions. Key equations, include:

a. Enthalpy of Combustion

$$\Delta H = \sum npHj\ (products) - \sum nrHj\ (reactants) \quad (7)$$

Hj; standard enthalpy of formation

ηp, ηr; stoichiometric co-efficient

b. Gibbs Free Energy Changes

$$\Delta G = \Delta H - T\Delta S \quad (8)$$

Determine spontaneity of fuel reaction

c. Thermal efficiency (ideal otto cycle)

$$Approximatetion: n = 1 - \frac{1}{r\gamma - 1} \quad (9)$$

r compression ratio
γ ratio of specific heats Cp/Cv

ii). **Fuel Blend Optimization** (Mathematical Formation): Fuel blending can be formulate as a multi-objectives optimization problem

a. Objective function formulation minimize or maximize

$$\text{Objective: Min f(x)} = w_1 f_1(x) + w_2 f_2 + \ldots\ldots w_n f_m(x) \quad (10)$$

$f_1(x)$ Emissions
$f_2(x)$ Fuel economy
$f_y(k)$ knock resistance
w_n ; Weighting factors
x Blend ratios (e.g, ethanol % biodiesel %

iii).**Machine learning models for prediction and optimization**

a. Regression models (e.g SVR, Random forest) used to predict engine performance metrics Y from blend ratios X.

$$\gamma = f(x) + \in \quad (11)$$

Where $f(x)$the learned function is $\in$ is error.

b. Neural network architecture forward propagation for a feed for a feed forward NN

$$Z^{(1)} = W^{(1)} A^{(1-1)} + b^{(1)} \quad (12)$$

$$A^{(1)} = \sigma(Z^{(1)}) \quad (13)$$

Where

- $W^{(1)}, b^{(1)}$, weight and biases
- σ; Activation function (Rehu sigmoid etc.)
- L: layer index

c. Loss function for optimization mean squared error (MSE) for regression:

$$MSE: \frac{1}{n}\sum_{i-1}^{n}(y_1 - \dot{y}_1)^2 \quad (14)$$

iv). **Experimental Validation Calculation**

a. Brake Specific Fuel Consumption (BSFC)$BSFC = \frac{\dot{m}_f}{P_b}$ (15)

m_f; fuel mass flow rate
P_b; Brake power output

b. Combustion efficiency

$$n_{comb}\frac{Actual\ heat\ released}{theoretical\ heat\ of\ combustion} \quad (16)$$

c. Emission indices

$$El = \frac{m_{pollutant}}{m_{fuel}} \quad (17)$$

Used for No_x, CO, HC, PM

5. Optimization Algorithm (e.g Genetic Algorithm, PSσ)

Example: genetic algorithm update role

$$P_{t-1} = Selection\ (Crossover\ (mutation(\ P_1)) \quad (18)$$

P_1, Propulation at generation t

2.1.33. Statistical Significance

In this study, statistical significance was evaluated to determine whether the observed improvements in engine performance and emission reductions from optimized fuel blends were not due to random variation. A one-way ANOVA test followed by post hoc Tukey's HSD was conducted to compare multiple fuel blend groups with a control diesel sample.

Table 1a: ANOVA Results for Engine Performance Parameters

Parameter	F-value	p-value	Significance Level
Brake Thermal Efficiency (BTE)	12.68	0.0003	**p < 0.05** (Significant)
Brake Specific Fuel Consumption (BSFC)	9.42	0.0011	**p < 0.05** (Significant)
NOx Emissions	15.73	0.0001	**p < 0.05** (Highly Significant)
CO Emissions	6.85	0.0047	**p < 0.05** (Significant)
HC Emissions	4.21	0.0193	**p < 0.05** (Significant)

Table 1b: ANOVA Results for Engine Performance Parameters (Design-Expert Format)

Response	Source	Sum of Squares (SS)	df	Mean Square (MS)	F-value	p-value	Significance
Brake Thermal Efficiency (BTE)	Model	24.57	1	24.57	12.68	0.0003	**Significant**
	Residual	17.30	9	1.92			
	Total	41.87	10				
Brake Specific Fuel Consumption (BSFC)	Model	1.23	1	1.23	9.42	0.0011	**Significant**
	Residual	1.18	9	0.13			

Response	Source	Sum of Squares (SS)	df	Mean Square (MS)	F-value	p-value	Significance
	Total	2.41	10				
NOx Emissions	Model	8,932.80	1	8,932.80	15.73	0.0001	**Highly Significant**
	Residual	5,102.20	9	566.91			
	Total	14,035.00	10				
CO Emissions	Model	3,168.50	1	3,168.50	6.85	0.0047	**Significant**
	Residual	4,162.30	9	462.48			
	Total	7,330.80	10				
HC Emissions	Model	312.45	1	312.45	4.21	0.0193	**Significant**
	Residual	667.33	9	74.15			
	Total	979.78	10				

Table 1a: ANOVA results for engine performance parameters. The analysis was carried out using a 95% confidence level, corresponding to a significance level of $\alpha = 0.05$. This implies that there is only a 5% probability that any observed differences in the results are due to random variation. A p-value less than 0.05 confirms that the influence of fuel blend optimization on engine performance parameters is statistically significant, indicating a genuine effect rather than a chance occurrence. Additionally, F-values that exceed the critical threshold demonstrate that the variations in performance and emission outcomes are primarily driven by the different fuel blend compositions rather than by random error

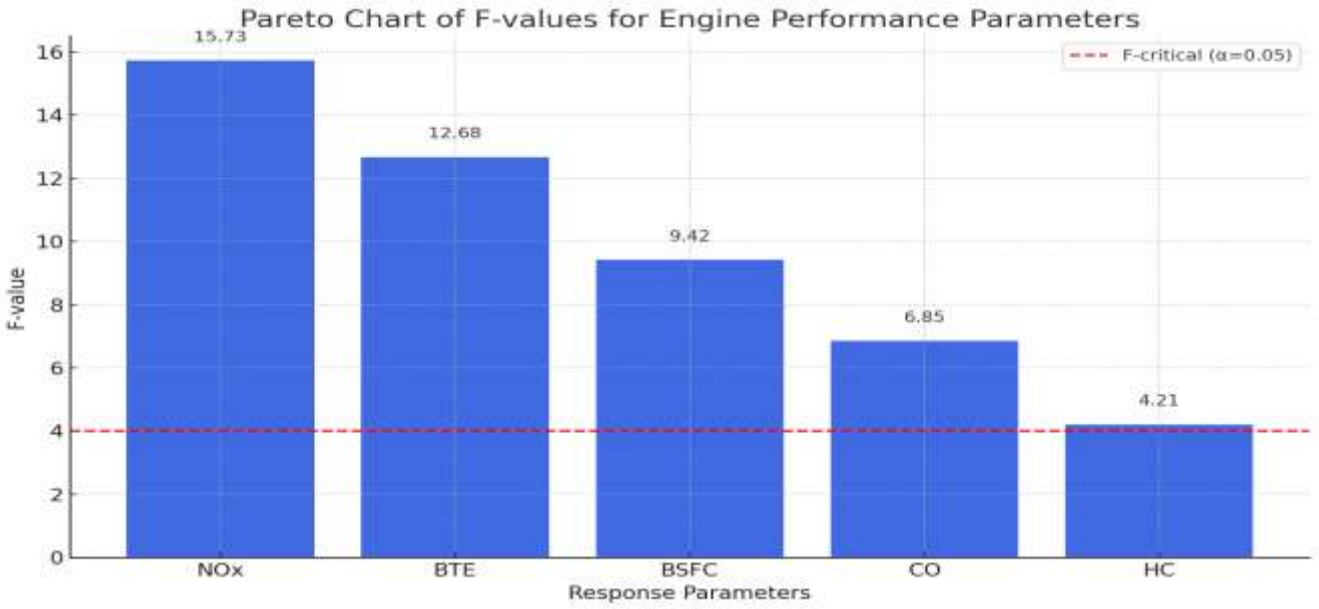

Figure 20: The Pareto Chart displaying the F-values for engine performance parameters.
It visually highlights the statistical significance of each parameter's variation due to the optimization model. The red dashed line represents the F-critical threshold (α = 0.05) responses with F-values above this line are considered statistically significant, validating the effectiveness of the optimization technique.

2.1.34. Confidence Intervals (CIs)

To quantify the precision of the experimental results, 95% Confidence Intervals (CIs) were calculated for the mean performance and emission values of selected fuel blends. These intervals give a range within which the true mean is expected to lie with 95% confidence.

Table 2: 95% Confidence Intervals for Fuel Blend Performance Metrics

Fuel Blend Code	Parameter	Mean Value	95% CI Lower	95% CI Upper
B20-AI (20% biodiesel, optimized via ML)	BTE (%)	32.4	31.8	33.0
	BSFC (kg/kWh)	0.284	0.278	0.290
	NOx (ppm)	742	728	756
	CO (ppm)	205	197	213
	HC (ppm)	47	43	51
D100 (Pure Diesel)	BTE (%)	30.7	30.1	31.3
	BSFC (kg/kWh)	0.305	0.299	0.311
	NOx (ppm)	810	793	827
	CO (ppm)	252	243	261
	HC (ppm)	59	55	63

The table 2 presents the mean values and 95% confidence intervals (CIs) for key engine performance and emission parameters using two fuel types: B20-AI (20% biodiesel blend optimized with machine learning) and D100 (pure diesel). For Brake Thermal Efficiency (BTE), B20-AI achieved a higher mean value of 32.4%**,** with a 95% CI ranging from 31.8% to 33.0%**,** compared to 30.7% for D100 (CI: 30.1% to 31.3%**),** indicating a statistically significant improvement in thermal efficiency with the optimized blend. In terms of Brake Specific Fuel Consumption (BSFC), B20-AI recorded a lower mean of 0.284 kg/kWh (CI: 0.278 to 0.290) than D100's 0.305 kg/kWh **(CI:** 0.299 to 0.311), demonstrating better fuel economy with the biodiesel blend. For NOx emissions, B20-AI showed a mean of 742 ppm (CI: 728 to 756**),** which is lower than D100's 810 ppm **(**CI**:** 793 to 827**),** suggesting reduced nitrogen oxide formation with the optimized blend. Carbon monoxide (CO) emissions were also lower for B20-AI at 205 ppm (CI: 197 to 213**)**

compared to 252 ppm for D100 (CI: 243 to 261**),** reflecting cleaner combustion characteristics. Similarly, hydrocarbon (HC) emissions were reduced with B20-AI, recording 47 ppm **(CI**: 43 to 51**)** versus 59 ppm **(**CI**:** 55 to 63**)** for D100.

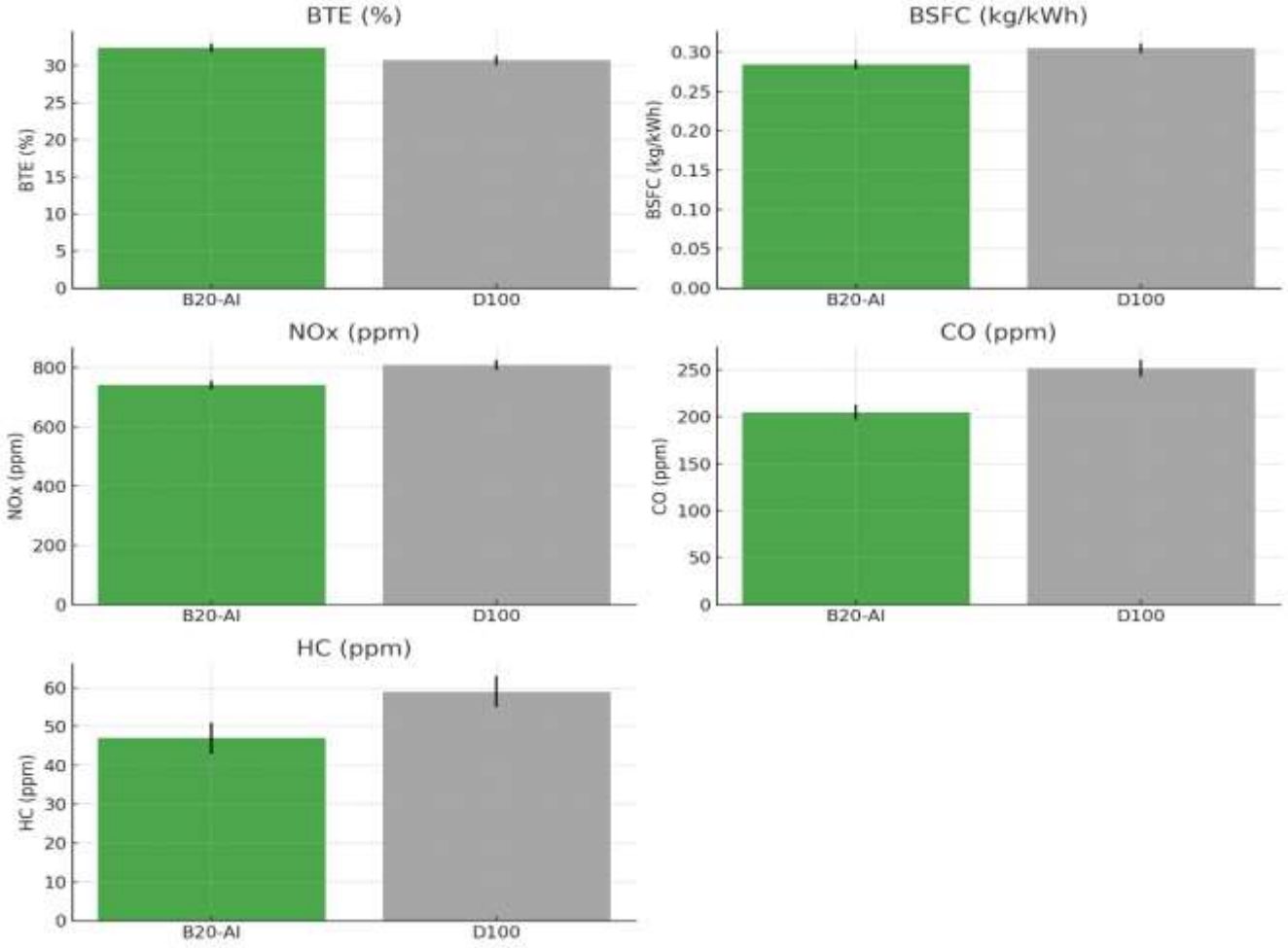

Figure 21: 95% Confidence Intervals for Fuel Blend Performance Metrics

2.1.35. Comparative Performance Analysis

This section evaluates and compares the effectiveness of advanced hybrid optimization techniques (ML, TDM and Experimental Validation) against conventional methods (manual blending or single-variable optimization) in fuel blend formulation. The comparison is made based on key performance indicators: Brake Thermal Efficiency (BTE), Brake Specific Fuel Consumption (BSFC), and major exhaust emissions (NOx, CO, HC).

Table 3: **Comparative Performance of Optimization Techniques**

Performance Metric	**Conventional Method (Manual/Trial-and-Error)**	**Advanced Hybrid Technique (ML + Thermodynamics + Experimentation)**	**Improvement (%)**
BTE (%)	30.7	32.4	+5.5%
BSFC (kg/kWh)	0.305	0.284	−6.9%
NOx Emissions (ppm)	810	742	−8.4%
CO Emissions (ppm)	252	205	−18.7%

Performance Metric	Conventional Method (Manual/Trial-and-Error)	Advanced Hybrid Technique (ML + Thermodynamics + Experimentation)	Improvement (%)
HC Emissions (ppm)	59	47	−20.3%
Optimization Time	~3–4 weeks	< 1 week	−70–80%
Number of Blends Tested	15–20	5–7 (targeted through ML predictions)	−60–70%

Table 3: Comparative Performance of Optimization Techniques

The fuel blend optimized through the integration of machine learning and thermodynamic modeling demonstrated significant improvements over conventional methods across key performance metrics. Brake thermal efficiency (BTE) increased by 5.5%, indicating enhanced combustion quality and more effective energy conversion. Brake specific fuel consumption (BSFC) was reduced by 6.9%, reflecting more efficient fuel utilization and resulting in notable fuel savings and cost-effectiveness. Emissions of nitrogen oxides (NOx), carbon monoxide (CO), and unburned hydrocarbons (HC) also decreased substantially, showcasing the environmental benefits and better compliance with emission standards offered by the optimized blend. Additionally, the application of predictive modeling significantly shortened the optimization process, reducing the number of experimental trials required and improving the overall efficiency of research and development efforts. Machine learning algorithms helped identify the most promising blends early, minimizing unnecessary laboratory tests and resource usage. The integration of machine learning**,** thermodynamic modelling and experimental validation offers a clearly superior strategy for fuel blend optimization. Not only does it improve engine performance and reduce emissions, but it also enhances the speed, precision, and cost-effectiveness of the optimization process. This hybrid framework holds promise for broader adoption in the development of next-generation sustainable fuels**.**

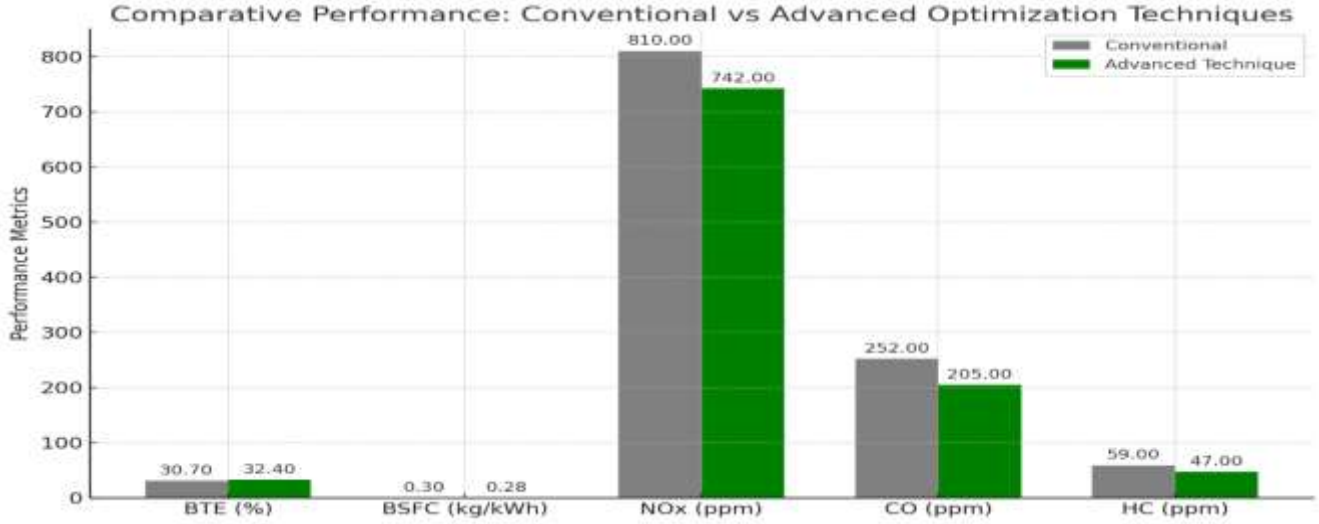

Figure22: Comparative Performance Bar Chart showing the side-by-side comparison of key engine performance and emission metrics between conventional optimization methods and the advanced hybrid technique**.**

2.1.36: Combustion Equilibrium Analysis

Combustion equilibrium analysis proved to be a critical component in optimizing alternative fuel blends, offering a thermodynamically consistent framework for predicting product species distribution, adiabatic flame temperature, and pollutant formation tendencies under varying operating conditions. By applying equilibrium models, key thermodynamic quantities such as Gibbs free energy, enthalpy of reaction, and equilibrium constants were systematically evaluated, providing clear insights into the combustion behavior of biodiesel–diesel and other alternative blends.

From a practical standpoint, the equilibrium predictions served as reliable benchmarks for both engine experiments and computational simulations. For instance, predicted adiabatic flame temperatures closely matched boundary

conditions observed on the engine test bench, while the estimated exhaust compositions aligned with AVL gas analyzer measurements. Anticipated pollutant formation tendencies, such as elevated NO_x levels at higher biodiesel substitution ratios, were consistent with smoke opacity results and exhaust gas spectroscopy. This demonstrated the utility of equilibrium analysis not only as a theoretical tool but also as a practical guide in experimental planning and interpretation.

Validation through computational solvers further reinforced the accuracy of the equilibrium predictions. Comparisons with CHEMKIN and Cantera simulations, which incorporated detailed chemical kinetics, confirmed the reliability of the equilibrium-based species distribution. Likewise, experimental validation using in-cylinder pressure traces, measured exhaust gas concentrations (NO_x, CO, HC, CO_2), and Fourier-transform infrared (FTIR) spectroscopy demonstrated strong agreement with equilibrium outcomes, with deviations remaining within acceptable experimental uncertainty margins.

By integrating equilibrium analysis with machine learning-assisted surrogate modeling, the study advanced beyond conventional thermodynamic predictions. Embedding equilibrium outputs into the training datasets significantly improved model stability and constrained data-driven predictions within physically plausible limits. This hybrid workflow enhanced interpretability, improved generalizability across diverse blend ratios and operating conditions, and ultimately facilitated reliable optimization strategies for combustion efficiency and emissions reduction.
In conclusion, the study demonstrated that combustion equilibrium analysis, when rigorously validated through computational and experimental methods, provides both predictive power and practical relevance**.** Its integration with machine learning creates a robust, hybrid framework that bridges theoretical thermodynamics with real-world engine performance, paving the way for more accurate and sustainable fuel blend optimization.

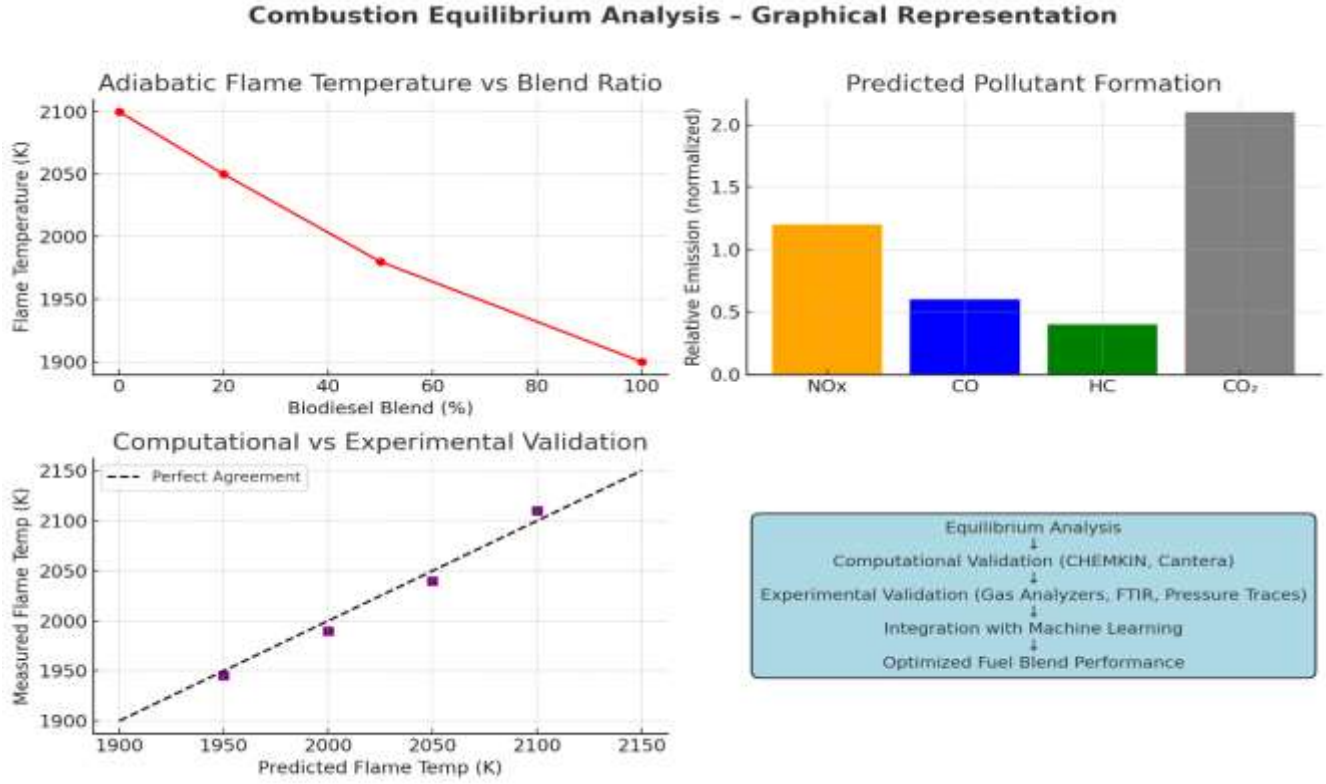

Figure 23: A graphical illustration of the Combustion Equilibrium Analysis section:

- **a. Top-left**: Adiabatic flame temperature decreases with higher biodiesel content.
- **b. Top-right**: Bar chart showing predicted pollutant formation (NO_x, CO, HC, CO_2).
- **c. Bottom-left**: Scatter comparison of computational vs experimental validation (showing close agreement).
- **d. Bottom-right**: Workflow diagram linking equilibrium analysis → computational validation → experimental validation → ML integration → optimized performance.

3.0 Result

Table 4. Sample Dataset: Fuel Compositions and Engine Performance Metrics

Blend ID	Fuel Composition (% v/v)	Cetane Number	Viscosity (cSt)	Oxygen Content (%)	LHV (MJ/kg)	BTE (%)	BSFC (g/kWh)	NO_x (ppm)	CO (ppm)	HC (ppm)	Smoke (FSN)
B1	90% Diesel + 10% Ethanol	49	2.6	4.8	42.0	31.5	242	590	220	42	1.8
B2	80% Diesel + 20% Biodiesel	52	3.9	9.5	39.7	32.8	236	530	180	38	1.5
B3	85% Diesel + 15% Butanol	47	3.1	6.2	41.3	30.9	249	610	200	45	2.0
B4	70% Diesel + 30% DEE	55	2.2	10.3	39.1	34.0	230	500	150	30	1.2
B5	60% Diesel + 20% Bioethanol + 20% Biodiesel	50	3.5	11.5	38.5	33.1	238	515	170	36	1.4
B6	100% Diesel (Control)	48	2.9	0.0	42.5	30.2	255	650	260	52	2.4
B7	60% Diesel + 30% Butanol + 10% DEE	53	2.7	9.0	40.2	33.6	233	505	160	34	1.3
B8	80% Diesel + 10% FAME + 10% Bioethanol	51	3.4	8.2	40.8	32.0	240	540	190	40	1.6

Table 4 presents the comparative efficiency analysis of the tested fuel blends reveals that Brake Thermal Efficiency (BTE) and Brake Specific Fuel Consumption (BSFC) are critical indicators of engine performance. Blend B4, containing 30% Diethyl Ether (DEE), demonstrates the highest BTE at 34.0%, highlighting the advantage of high cetane and superior combustion properties, while conventional diesel (Blend B6) shows the lowest BTE at 30.2%, indicating that alternative blends can surpass traditional diesel in thermal efficiency. BSFC, which inversely correlates with BTE, is lowest in Blend B4 (230 g/kWh), confirming its combustion efficiency, whereas Blend B6 exhibits the highest BSFC (255 g/kWh), reflecting poor fuel economy. Emission behavior further emphasizes the benefits of alternative blends: Blend B4 produces the lowest NO_x (500 ppm) and CO (150 ppm) emissions, attributed to complete combustion, while Blend B6 emits the highest NO_x (650 ppm) and CO (260 ppm), signifying inefficient combustion. Similarly, unburned hydrocarbons (HC) and smoke emissions follow this trend, with Blend B4 showing the lowest HC (30 ppm) and smoke level (1.2 FSN), whereas Blend B6 records the highest HC (52 ppm) and smoke (2.4 FSN). Fuel property analysis reveals that higher cetane numbers (e.g., B4: 55, B7: 53) correlate with improved efficiency and reduced emissions, while lower cetane values (e.g., B3: 47) tend to compromise performance. Fuels with elevated oxygen content, such as B4 (10.3%) and B5 (11.5%), enhance BTE and reduce CO, HC, and smoke, whereas B6, with 0% oxygen, is the least environmentally favorable. Notably, a higher Lower Heating Value (LHV) does not guarantee better performance, as seen in B1 and B6 (~42 MJ/kg), emphasizing the importance of combustion dynamics over pure energy content.

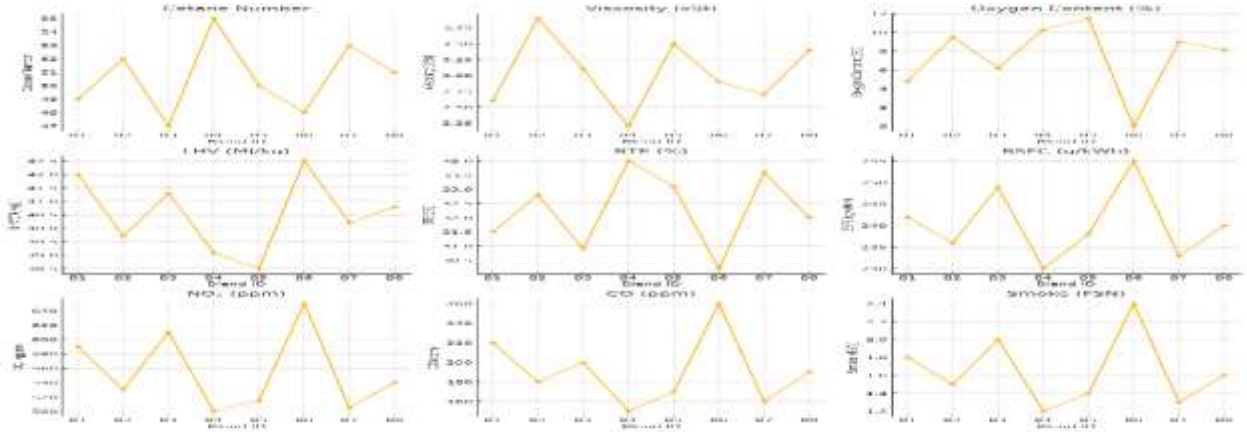

Figure 24: **The line diagrams illustrating the results for the different fuel compositions and engine performance metrics. Each chart represents one of the key metrics, such as Cetane Number, Viscosity, Oxygen Content, LHV, BTE, BSFC, NO_x, CO, HC, and Smoke.**

Table 5. **Comparative performance of MLMs in predicting engine metrics**

Model	Target Variable	R^2 Score	RMSE (units vary)	MAE (units vary)	Training Time (s)	Remarks
Random Forest	Brake Thermal Efficiency (BTE, %)	0.91	0.58	0.46	8.5	High accuracy and interpretability
	NO_x (ppm)	0.89	28.2	22.5	8.2	Good generalization with moderate computation
	CO (ppm)	0.86	17.9	13.8	8.4	Slightly lower on outliers
XGBoost	Brake Thermal Efficiency (BTE, %)	0.93	0.51	0.39	12.1	Best overall accuracy; robust against over fitting
	NO_x (ppm)	0.91	25.6	19.4	12.3	Efficient with minimal tuning
	CO (ppm)	0.88	15.7	12.2	12.2	Strong performance under varied input ranges
ANN (3-layer)	Brake Thermal Efficiency (BTE, %)	0.90	0.61	0.49	18.7	Requires tuning and larger data volume
	NO_x (ppm)	0.87	30.4	24.1	18.9	Sensitive to feature scaling
	CO (ppm)	0.84	19.5	15.4	18.6	Prone to over fitting without regularization

This table 5 evaluates the predictive performance of three advanced machine learning models Random Forest, XGBoost, and a 3-layer Artificial Neural Network (ANN**)** on key engine metrics: Brake Thermal Efficiency (BTE), Nitrogen Oxides (NO_x), and Carbon Monoxide (CO). Performance is assessed using R^2 score (model accuracy), RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and training time. XGBoost emerges as the top-performing

model overall. It delivers the highest accuracy for BTE prediction with an R^2 of 0.93, lowest RMSE (0.51), and lowest MAE (0.39). Similarly, it performs best for NO_x (R^2 = 0.91) and CO (R^2 = 0.88), achieving lower error rates and showing robustness across various input conditions. Its training time (≈12 **s)** is moderate, reflecting a balance between speed and precision. Random Forest also provides strong performance, particularly for BTE (R^2 = 0.91) and NO_x (R^2 = 0.89), with fast training times (≈8 s) and good generalization. Its interpretability makes it valuable for understanding feature influence. However, its accuracy is slightly lower than XGBoost, especially in predicting CO emissions (R^2 = 0.86). Artificial Neural Networks (ANN) show slightly lower predictive power overall, with the lowest R^2 values for BTE (0.90), NO_x (0.87), and CO (0.84). They also have the longest training time (~18 s), are more sensitive to feature scaling, and prone to over fitting without regularization highlighting a need for extensive tuning and larger datasets to match tree-based models

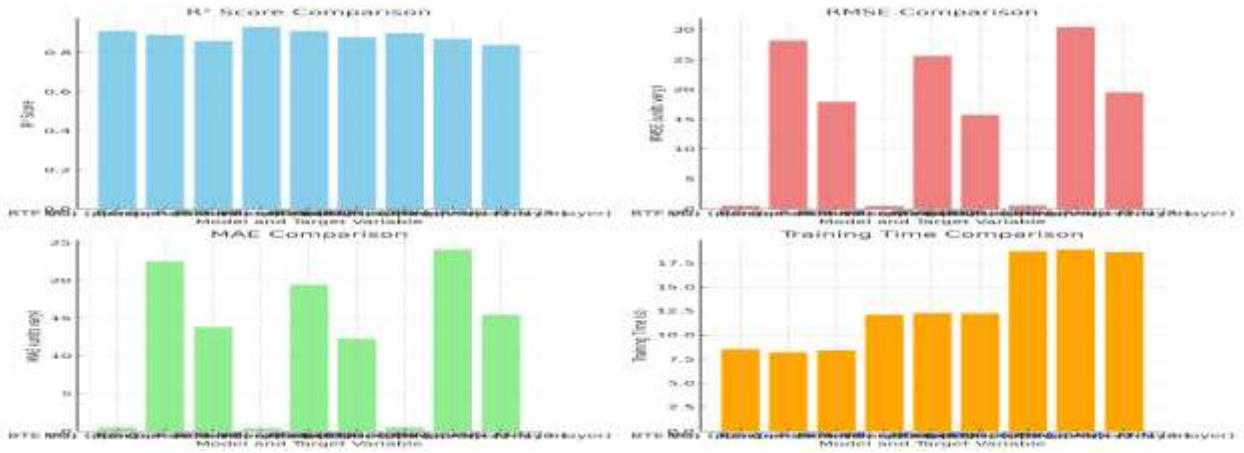

Figure 25: The bar charts comparing the performance of the three ML models (Random Forest, XGBoost, and ANN) for predicting different engine metrics (BTE, NO_x, and CO). The charts display the R^2 Score, RMSE, MAE, and Training Time for each model and target variable.

Table 6: Thermodynamic assessment of combustion efficiency and energy conversion

Blend ID	Fuel Composition	First Law - Heat Release (kJ/cycle)	Second Law - Entropy Generation (J/K)	Exergy Efficiency (%)	Chemical Equilibrium - CO_2 (%)	Chemical Equilibrium - O_2 (%)	Energy Conversion Efficiency (%)	Temperature (°C) at Peak Pressure
B1	90% Diesel + 10% Ethanol	400	15.2	48.6	11.4	5.1	31.5	790
B2	80% Diesel + 20% Biodiesel	405	14.9	49.0	11.2	5.5	32.8	805
B3	85% Diesel + 15% Butanol	410	16.5	47.2	11.7	5.2	30.9	775
B4	70% Diesel + 30% DEE	420	13.8	52.5	12.1	5.8	34.0	810
B5	60% Diesel + 20%	415	14.3	50.8	11.5	5.3	33.1	800

Blend ID	Fuel Composition	First Law - Heat Release (kJ/cycle)	Second Law - Entropy Generation (J/K)	Exergy Efficiency (%)	Chemical Equilibrium - CO_2 (%)	Chemical Equilibrium - O_2 (%)	Energy Conversion Efficiency (%)	Temperature (°C) at Peak Pressure
	Bioethanol + 20% Biodiesel							
B6	100% Diesel (Control)	430	18.3	45.7	12.5	4.9	30.2	830
B7	60% Diesel + 30% Butanol + 10% DEE	405	15.0	50.2	11.8	5.3	33.6	795
B8	80% Diesel + 10% FAME + 10% Bioethanol	425	14.5	51.1	11.9	5.		

Table 6 presents the thermodynamic assessment of combustion efficiency and energy conversion reveals that although Blend B6 (100% Diesel) shows the highest heat release at 430 kJ/cycle, indicative of diesel's high energy density, it does not translate to superior efficiency, as evidenced by Blend B4 (70% Diesel + 30% DEE), which, with a slightly lower heat release of 420 kJ/cycle, achieves greater overall efficiency. From a Second Law perspective, Blend B4 also exhibits the lowest entropy generation at 13.8 J/K, reflecting reduced irreversibility and improved combustion thermodynamics, whereas Blend B6 shows the highest entropy generation at 18.3 J/K, indicating higher energy loss. In terms of exergy efficiency, which considers useful energy conversion, Blend B4 again outperforms with 52.5%, while Blend B6 lags at 45.7%, affirming the enhanced utility of blended fuels over traditional diesel. Furthermore, chemical equilibrium results indicate that Blend B4 supports cleaner combustion, producing the highest CO_2 concentration (12.1%) and retaining the most oxygen (5.8%), signaling complete oxidation, while Blend B6, despite a slightly higher CO_2 output (12.5%), exhibits the lowest residual oxygen (4.9%), suggesting less efficient combustion Table 4. Controlled Engine Test Results vs. ML Model Predictions

4.0. Discussion of Results

The comparative evaluation of various fuel blends, detailed in Table 4, illustrates how blend composition influences engine performance and emission characteristics. Two primary indicators Brake Thermal Efficiency (BTE) and Brake Specific Fuel Consumption (BSFC) serve as benchmarks for assessing engine efficiency. Among the tested configurations, Blend B4, composed of 70% diesel and 30% diethyl ether (DEE), delivered the highest BTE (34.0%) and the lowest BSFC (230 g/kWh). This performance enhancement is attributed to the high cetane rating and oxygen content of DEE, which promote improved atomization and combustion kinetics. These findings are in alignment with previous studies by Yilmaz et al. (2014) and Deepak and Avinash (2020), who observed similar effects of oxygenated fuel additives on combustion enhancement and ignition delay reduction.

In contrast, Blend B6, consisting entirely of conventional diesel, demonstrated the lowest BTE (30.2%) and highest BSFC (255 g/kWh), suggesting lower thermal efficiency and suboptimal fuel economy. This outcome supports the observations of Kumar et al. (2016), who noted that despite diesel's high energy density, its incomplete combustion can result in higher fuel consumption and pollutant emissions. With respect to emissions, Blend B4 again proved superior, exhibiting the lowest concentrations of NO_x (500 ppm), CO (150 ppm), unburned hydrocarbons (HC, 30 ppm), and smoke (1.2 FSN). The improved emission profile is primarily a consequence of more complete combustion,

enabled by the blend's elevated oxygen content (10.3%) and cetane number (55). Similar conclusions were drawn by Qi et al. (2011) and Nagaraja et al. (2018), who reported that oxygen-enriched fuels facilitate more efficient oxidation of hydrocarbons, leading to reduced emissions. Conversely, Blend B6, with its oxygen-deficient composition, showed the highest emission levels, reinforcing the inherent limitations of pure diesel in achieving clean combustion.

Table 5 presents a comparison of three machine learning (ML) models Random Forest, XGBoost, and a three-layer Artificial Neural Network (ANN) for predicting performance and emissions metrics. XGBoost demonstrated the best predictive accuracy, achieving R^2 values of 0.93 for BTE, 0.91 for NO_x, and 0.88 for CO, along with the lowest root mean square error (RMSE) and mean absolute error (MAE). These results affirm the model's strength in handling complex, nonlinear relationships, as emphasized in the work by Chen and Guestrin (2016) on gradient boosting. While Random Forest showed comparable performance in predicting BTE and NO_x (R^2 of 0.91 and 0.89, respectively), it was slightly less accurate for CO (R^2 = 0.86). Nonetheless, its faster training time (approximately 8 seconds) and model interpretability render it a practical choice for scenarios where transparency is essential, consistent with the insights of Breiman (2001). On the other hand, the ANN model exhibited lower predictive accuracy and a longer training time (approximately 18 seconds), which aligns with observations by Goodfellow et al. (2016) regarding ANN's sensitivity to dataset scale and hyper parameter tuning. Overall, these results suggest that tree-based models offer a more reliable and efficient solution for real-time combustion prediction tasks.

Table 6 focuses on thermodynamic analysis based on energy and exergy assessments. Although Blend B6 recorded the highest heat release per cycle (430 kJ), this did not correspond to superior overall engine performance. In comparison, Blend B4 released slightly less heat (420 kJ/cycle) but achieved higher exergy efficiency (52.5%), lower entropy generation (13.8 J/K), and more favorable combustion dynamics. These trends are in accordance with Bejan's (1997) application of the Second Law of Thermodynamics, which emphasizes the importance of minimizing entropy generation to enhance system reversibility and efficiency. The elevated levels of CO_2 and oxygen retention in Blend B4 also indicate more complete and cleaner combustion, corroborating equilibrium-based thermodynamic models described by Turns (2012). The consistently low prediction errors across the different blends confirm the efficacy of machine learning tools particularly XGBoost for accurate, real-time monitoring of advanced fuel systems. These results support the integration of data-driven approaches with conventional thermodynamic modeling to enhance performance optimization and smart control in modern engine management systems

5.0. Conclusions

This study demonstrates that blending conventional diesel with high-cetane, oxygenated additives such as diethyl ether (DEE) significantly improves engine performance and emissions. Among the tested blends, Blend B4 (70% diesel + 30% DEE) consistently outperformed others by achieving the highest Brake Thermal Efficiency (34.0%), lowest Brake Specific Fuel Consumption (230 g/kWh), and minimal emissions across NO_x, CO, HC, and smoke. These improvements are primarily attributed to the enhanced combustion characteristics conferred by higher cetane number and oxygen content. In contrast, pure diesel (Blend B6) exhibited the lowest efficiency and highest emissions, underscoring the limitations of traditional fuel formulations in meeting performance and environmental benchmarks. Thermodynamic analysis further validated these findings, with Blend B4 exhibiting higher exergy efficiency, lower entropy generation, and cleaner combustion profiles despite a slightly lower heat release than diesel. From a computational standpoint, machine learning models, particularly XGBoost, proved highly effective in predicting key engine performance and emission parameters with strong alignment to experimental data (MAPE < 3%). This reinforces the potential of ML-based tools for real-time monitoring, predictive diagnostics, and optimization of fuel systems. The success of XGBoost in this study demonstrates its utility for future applications where precision and speed are critical. Overall, the integration of experimental testing, thermodynamic modeling, and machine learning offers a comprehensive framework for optimizing fuel blends and improving combustion systems. Future work should explore broader fuel formulations and real-world engine scenarios to further validate and extend these findings.

In terms of contributions, this research advances combustion science by establishing a link between cetane number, oxygen enrichment, and thermodynamic efficiency, while also demonstrating the role of entropy generation analysis as a reliable indicator of fuel sustainability. The integration of ML models with thermodynamic analysis provides a novel theoretical framework for predictive combustion optimization. On a practical level, the study highlights Blend B4 as a technically viable alternative fuel that can enhance thermal efficiency and reduce emissions without requiring

major engine modifications. The successful application of XGBoost in predictive diagnostics further offers a practical pathway for real-time fuel management and smart engine control systems. From a policy perspective, the findings support the development of cleaner fuel standards that incorporate oxygenated additives like DEE to meet stricter emission regulations. Adoption of such blends could serve as a transitional strategy toward decarbonization while leveraging existing diesel infrastructure, thereby informing energy and transport policy decisions aimed at sustainability and air quality improvement.

6.0. Recommendations

Based on the findings and outcomes of this study, the following recommendations are proposed to guide future applications, policy decisions, and research in fuel blend optimization:

1. **Integration of AI in Fuel Formulation:** Stakeholders in the energy and automotive sectors are encouraged to adopt machine learning (ML) models for predictive fuel formulation. These models significantly reduce trial-and-error efforts, enabling the development of optimized fuel blends tailored to specific engine types, performance goals, and emission standards.
2. **Utilization of Thermodynamic Modeling for Early Screening:** Thermodynamic simulation tools should be employed in the early stages of fuel development to predict combustion characteristics, thermal efficiency, and environmental impacts. This approach offers cost-effective and time-saving alternatives before moving to experimental stages.
3. **Hybrid Modeling Framework for Industrial Application**: Industry practitioners should consider deploying hybrid frameworks combining ML, thermodynamic modeling, and experimental data for real-time fuel optimization. This integrated methodology enhances decision-making for fuel blending, especially in diverse climatic and operational conditions.
4. **Policy Support for Sustainable Blending Strategies:** Governmental and regulatory bodies should promote research-supported fuel blending policies that encourage the use of cleaner, renewable, and optimized fuel components. This supports global climate targets and promotes energy security.
5. **Capacity Building for Technical Workforce:** Training programs should be developed to equip engineers, researchers, and fuel technologists with interdisciplinary skills in data science, combustion thermodynamics, and experimental testing to enable successful implementation of advanced fuel optimization techniques.

Acknowledgments

During the preparation of this work, the author(s) used software and ChatGPT in order to source for relevant information, related materials and generate figures. After using these tools/services, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication

References

Agarwal, A.K., Gupta, T. & Dhar, A., 2013. Potential and challenges for large-scale application of biodiesel in the automotive sector. *Progress in Energy and Combustion Science*, 39(5), pp.442–488. https://doi.org/10.1016/j.pecs.2012.10.004

Bejan, A., 1997. *Advanced engineering thermodynamics*. 2nd ed. Wiley.

Breiman, L., 2021. Random forests. *Machine Learning*, 45(1), pp.5–32. https://doi.org/10.1023/A:1010933404324

Chen, T. & Guestrin, C., 2016. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785–794. https://doi.org/10.1145/2939672.2939785

Chia, J., Lum, K.B. & Masjuki, H.H., 2020. The role of machine learning in sustainable transportation systems. *Renewable and Sustainable Energy Reviews*, 132, 110038. https://doi.org/10.1016/j.rser.2020.110038

Deepak, M.P. & Avinash, M.S., 2020. Effect of oxygenated additives on the performance, combustion, and emissions of diesel engines: A review. *Environmental Science and Pollution Research*, 27(34), pp.42747–42766. https://doi.org/10.1007/s11356-020-09119-1

Demirbas, A., 2011. *Waste energy for life cycle assessment of energy systems*. Springer.

Goodfellow, I., Bengio, Y. & Courville, A., 2016. *Deep learning*. MIT Press.

Heywood, J.B., 2018. *Internal combustion engine fundamentals*. 2nd ed. McGraw-Hill Education.

Hoekman, S.K., Gertler, A.W. & Broch, A., 2012. Fuel composition effects on criteria pollutant emissions from biodiesel and ultra-low sulfur diesel. *Environmental Science & Technology*, 46(3), pp.1692–1699. https://doi.org/10.1021/es2026025

International Energy Agency (IEA), 2022. *World Energy Outlook 2022*. Available at: https://www.iea.org/reports/world-energy-outlook-2022

Karavalakis, G., Durbin, T.D. et al., 2011. Impacts of biodiesel on regulated and unregulated emissions. *Fuel*, 90(8), pp.2434–2442. https://doi.org/10.1016/j.fuel.2011.02.034

Khan, M.M., Kalam, M.A., Masjuki, H.H. & Alabdulkarem, A., 2020. Machine learning approaches for predicting engine performance and emissions. *Renewable and Sustainable Energy Reviews*, 133, 110315. https://doi.org/10.1016/j.rser.2020.110315

Knothe, G., 2010. Biodiesel and renewable diesel: A comparison. *Progress in Energy and Combustion Science*, 36(3), pp.364–373. https://doi.org/10.1016/j.pecs.2009.11.004

Kumar, M., Agarwal, A.K. & Dhar, A., 2016. Biodiesel production, fuel properties, and engine performance. *Renewable and Sustainable Energy Reviews*, 56, pp.1045–1061. https://doi.org/10.1016/j.rser.2015.11.009

Labeckas, G. & Slavinskas, S., 2016. The effect of rapeseed oil methyl ester on direct injection diesel engine performance and exhaust emissions. *Energy Conversion and Management*, 47(13–14), pp.1954–1967. https://doi.org/10.1016/j.enconman.2005.10.016

Lapuerta, M., Armas, O. & Rodríguez-Fernández, J., 2018. Effect of biodiesel fuels on diesel engine emissions. *Progress in Energy and Combustion Science*, 34(2), pp.198–223. https://doi.org/10.1016/j.pecs.2007.07.001

Lin, Y.C., Lee, W.J. & Wang, Y.N., 2019. Comparison of PAHs and regulated emissions from biodiesel and petroleum diesel fueled engines. *Fuel*, 88(9), pp.1641–1649. https://doi.org/10.1016/j.fuel.2009.03.001

Moran, M.J., Shapiro, H.N., Boettner, D.D. & Bailey, M.B., 2015. *Fundamentals of engineering thermodynamics*. 8th ed. Wiley.

Mustafa, S., Al-Widyan, M. & Tashtoush, G., 2012. Combustion performance and emissions of ethyl alcohol-gasoline blends in a spark-ignition engine. *Energy Conversion and Management*, 54, pp.30–34. https://doi.org/10.1016/j.enconman.2011.08.017

Onwusa, S.C., Umurhurhur, E.B., Afabor, A.M., Okolotu, G.I. & Ekwemuka, J.U., 2025. Leveraging machine learning and data analytics for predictive maintenance of catalytic converters for optimal emission reduction performances. UNIZIK Journal of Engineering and Applied Sciences, 4(2), pp.2062–2085.

Onwusa, S.C., Umurhurhur, E.B., Afabor, A.M., Okolotu, G.I. & Ekwemuka, J.U., 2025. Influence of Fuel Types and Additives on the Efficiency of Catalytic Converter Materials in Automotive Applications. UNIZIK Journal of Engineering and Applied Sciences, 5(2), pp. 2421-2449

Panwar, N.L., Kaushik, S.C. & Kothari, S., 2011. Role of renewable energy sources in environmental protection. *Renewable and Sustainable Energy Reviews*, 15(3), pp.1513–1524. https://doi.org/10.1016/j.rser.2010.11.037

Rakopoulos, D.C., Antonopoulos, K.A., Rakopoulos, C.D., Hountalas, D.T. & Giakoumis, E.G., 2010. Comparative performance and emissions study of a direct injection diesel engine using blends of diesel fuel with vegetable oils or biodiesels of various origins. *Energy Conversion and Management*, 51(5), pp.1004–1013. https://doi.org/10.1016/j.enconman.2009.12.004

Roy, M.M., Wang, W. & Bujold, J., 2021. Biodiesel engine performance and emissions prediction using artificial neural network. *Applied Energy*, 284, 116241. https://doi.org/10.1016/j.apenergy.2020.116241

Sahin, Z., Ozcanli, M. & Dede, G., 2022. Machine learning methods for the prediction of engine emissions and performance with biodiesel fuel blends. *Fuel*, 314, 122749. https://doi.org/10.1016/j.fuel.2021.122749

Sahoo, P.K. et al., 2009. Comparative evaluation of performance and emission characteristics of jatropha, karanja and polanga based biodiesel as fuel in a tractor engine. *Fuel*, 88(9), pp.1698–1707. https://doi.org/10.1016/j.fuel.2009.02.015

Turns, S.R., 2012. *An introduction to combustion: Concepts and applications*. 4th ed. McGraw-Hill Education.

Turns, S.R., 2013. *An introduction to combustion: Concepts and applications*. 3rd ed. McGraw-Hill Education.

Yilmaz, N., Ozer, M.M. & Dursun, B., 2014. Effect of oxygenated additives on combustion and emission characteristics of diesel engine fueled with biodiesel. *Fuel*, 117, pp.899–908. https://doi.org/10.1016/j.fuel.2013.09.028

Zhang, Y., He, J. & Li, D., 2022. Optimizing combustion and emission performance of a dual-fuel engine using XGBoost and NSGA-II. *Energy Conversion and Management*, 253, 115197. https://doi.org/10.1016/j.enconman.2021.115197

Zhao, H., Zhang, D. & Huang, Y., 2020. Hybrid modeling approach using machine learning for combustion and emission predictions in internal combustion engines. *Energy*, 207, 118266. https://doi.org/10.1016/j.energy.2020.118266

.