

AI MODELS FOR PREDICTING CONSTRUCTION DISPUTES IN SRI LANKA

Y.M.W.H.M.R.R.L.J.B. Kiridana¹, M.D.T.E. Abeynayake,² and B.A.I. Eranga³

ABSTRACT

Construction disputes pose persistent challenges in Sri Lanka's construction industry, leading to project delays, cost overruns, and strained professional relations. This research seeks to alleviate these issues by introducing an AI-powered predictive model designed to identify and analyse dispute risks at the project's outset. By offering proactive insights, the AI model aims to enhance decision-making and facilitate the implementation of dispute prevention strategies, thereby improving overall project outcomes. Employing a mixed-methods approach, the study comprehensively examined project features contributing to disputes within the Sri Lankan context. Quantitative data on project characteristics and their correlation with dispute occurrence were gathered through structured questionnaires, while qualitative insights into dispute causes and stakeholder challenges were obtained via in-depth interviews with industry experts. Through meticulous analysis of this combined data, key predictors of construction disputes were identified, including contract ambiguities, unrealistic timelines, payment delays, poor communication, and unforeseen site conditions. These findings drove the development of a machine learning-based predictive model trained to recognise patterns, predict dispute likelihoods, and suggest their nature based on identified risk factors. This innovative AI tool has the potential to revolutionise dispute management practices in Sri Lanka's construction industry. By providing stakeholders with early warnings of potential disputes, the model enables proactive mitigation strategies, such as enhanced contract drafting, optimised communication, and timely alternative dispute resolution. The long-term impact of this research extends to fostering a more collaborative and sustainable construction industry, ultimately contributing to the successful delivery of projects across Sri Lanka.

Keywords: *Artificial Intelligence; Causes of Construction Dispute; Construction Dispute; Construction Industry; Machine Learning.*

1. INTRODUCTION

The construction industry significantly contributes to economic development and overall quality of life. It enhances science and technology, strengthens infrastructure, and supports social and economic growth (Serogina et al., 2022). The construction sector stimulates growth, creates investment opportunities, and helps achieve national social and economic goals (Silva et al., 2018). Moreover, it plays a vital role in fulfilling the United

¹ Undergraduate, Department of Building Economics, University of Moratuwa, Sri Lanka, kiridanaymwhmrrlj.b.19@uom.lk

² Lecturer, Department of Building Economics, University of Moratuwa, Sri Lanka, mabeynayake@uom.lk

³ Lecturer, Department of Building Economics, University of Moratuwa, Sri Lanka, isurue@uom.lk

Nations' Sustainable Development Goals (SDGs) (Fei et al., 2021). Specifically, the industry is crucial for the SDGs related to equitable communities and cities, climate change mitigation, sanitation and water quality, responsible spending and manufacturing, and fostering innovation and infrastructure (Collier, 2020). Additionally, the construction industry is essential for the reconstruction of conflict-affected nations, contributing to employment, revenue generation, and the development of other industries (Serogina et al., 2022). Therefore, identifying common sources of disputes and improving the competency of judicial actors is crucial to making the resolution process more efficient (Marciano & Ramello, 2019). Factors such as risks, uncertainties, inadequate contract documentation, and behavioural issues contribute to the occurrence of construction disputes (Assaf et al., 2019). Moreover, analysing these factors and understanding their interrelationships is essential for effectively preventing and resolving disputes (Abeyasinghe & Jayathilaka, 2022).

The financial consequences of building conflicts are substantial. Cost overruns caused by litigation or delays can cripple projects, especially those with limited resources or those that serve important public infrastructure demands (Arfazadeh, 2014). Furthermore, the reputation of organisations involved in conflicts suffers, limiting their capacity to win future business (Ikechukwu et al., 2017). The hostile atmosphere created by conflicts hampers innovation and the development of collaborative working styles, which are increasingly viewed as threatening to the construction industry's long-term health (Walsh & Llp, 2007).

This research aims to equip stakeholders in the Sri Lankan construction industry with a powerful tool for anticipating and managing disputes effectively. By leveraging AI technology, inspiration is not only limited to mitigating the immediate impacts of conflicts yet fostering a more conducive environment for innovation, collaboration, and sustainable development. The creation and use of such technology have the potential to alter how the Sri Lankan construction sector predicts and resolves disputes. It enables project managers, engineers, and stakeholders to make data-driven decisions, promoting proactive dispute resolution (Choi et al., 2021). By anticipating disputes, AI technology might avoid costly escalations, maintain relationships, and ensure that projects are completed on time and within budget.

Finally, this has the potential to boost the overall efficiency, productivity, and reputation of Sri Lanka's construction industry (Sandagomika et al., 2020). Ultimately, by empowering decision-makers with data-driven insights and proactive strategies, the projects strive to contribute to the long-term prosperity and resilience of the construction sector in Sri Lanka, aligning with broader socio-economic objectives and the pursuit of equitable and sustainable growth.

2. LITERATURE REVIEW

2.1 DISPUTES IN CONSTRUCTION

Disputes in construction occur when conflicts between participants exist, and one party makes a claim which is rejected by the other (Yildizel et al., 2016). These disputes can arise from various issues such as contractual problems, delays, lack of communication, and design defects (El-Sayegh et al., 2020). Disputes in the construction industry have negative impacts on projects, leading to delays, loss of money, and occasionally project abandonment (Mashwama et al., 2019). While disputes in construction cannot be entirely

avoided, they can be minimised by ensuring that all project details are clearly stated and executed according to the agreed-upon terms (Osuzugbo & Okuntade, 2020). Therefore, a practical approach should be implemented to minimise disputes in the project (Gunarathna et al., 2018).

2.2 AI APPLICATION FOR DISPUTE PREDICTION

According to Agus et al. (2023), an AI-driven system has the potential to transform the prediction of disputes. Machine learning, as defined by El Naqa and Murphy (2015), is a rapidly evolving field of computing algorithms designed to mimic human intelligence by acquiring knowledge from the surrounding environment. Using machine learning, a model could analyse past project data, contractual agreements, communication patterns, and other relevant factors to identify key risk indicators (Agus et al., 2023). This information, presented through a user-friendly interface, enables project teams to monitor potential conflict triggers in real time (Choi et al., 2021). Importantly, the AI models can predict the likely severity of disagreements, helping stakeholders prioritise actions and efficiently allocate resources for dispute resolution (Chou et al., 2014). The development and implementation of such technology can transform how the Sri Lankan construction sector predicts and resolves disputes. It empowers project managers, engineers, and stakeholders to make data-driven decisions, promoting proactive dispute resolution (Choi et al., 2021). By anticipating disputes, AI technology can prevent costly escalations, and unhealthy relationships, and ensure the projects are completed on time and within the allocated budget. Eventually, this could enhance the overall efficiency, productivity, and reputation of Sri Lanka's construction industry (Sandagomika et al., 2020).

Machine learning algorithms designed to mimic human intelligence by acquiring knowledge from the surrounding environment operate as computational procedures, accomplishing tasks using input data without requiring explicit programming for predetermined results, as emphasised in the text (Horvitz & Mulligan, 2015). Unlike algorithms that are "hard-coded," these algorithms are "soft-coded," meaning they can alter and refine their structures based on repeated encounters or training (Molu & Goertz, 2014). During the training process, the input data is matched with the expected results, which enables the algorithm to fine-tune its configuration. This optimisation allows the algorithm to not only produce desirable outcomes using the training data, yet generalise and perform efficiently with the latest, undetected data (Regona et al., 2022). Training is the phase of machine learning that encompasses the process of "learning". Notably, the process of learning described above is not limited to a specific time frame. Like humans, a skilled algorithm can engage in continuous learning, referred to as "lifelong" learning, by consistently analysing fresh data and gaining knowledge from its errors (Liu, 2017). Supervised learning is considered a category of machine learning concept where the model is trained using labelled data in the training dataset, which means that the features which are selected for the training of the model are based on the labelled data (Zhou, 2018). Different algorithms generate functions that establish a correspondence between inputs and desired outcomes. An example involves classification tasks in which the algorithm is learning to estimate a function which maps input vectors to certain classes, identified as dependent on given input-output samples (Regona et al., 2022). Unsupervised learning is referred to as a type of machine learning where a model is trained by considering the unlabelled data without any specific guidance or supervision of the user. It is the process of representing a group of inputs without having any label or categorisation (Zhou, 2018). Furthermore, reinforcement learning is an algorithm that can

be used to learn how to make decisions by observing the environment and receiving feedback which can be used to guide the learning process, as the actions taken by the algorithm can lead to an impact on the environment. Learning to learn, the algorithm develops its own predefined biases based on past experiences (Caruana, 1993). In addition to these categories, machine learning algorithms are widely classified as supervised and unsupervised learning. In the context of supervised learning, predetermined categories are assigned to labelled data segments, which can usually be annotated by humans (Carcillo et al., 2021). The algorithm's main objective is to detect patterns and create mathematical models, which can be used to assess based on their predicted accuracy compared to measurements of data variation.

2.3 RESEARCH GAP

This research mainly focused on the construction industry of Sri Lanka because current approaches to dispute prediction in Sri Lanka frequently rely on subjective assessments and expert judgment (Gunarathna et al., 2018). While useful, such methods have limitations in terms of scalability, consistency, and the ability to analyse large amounts of data to find hidden patterns. Further, the traditional approaches may struggle to predict the cumulative impact of several risk factors throughout the various stages of a construction project (Wang et al., 2018). The lack of dependable early warning techniques limits stakeholders' ability to proactively minimise possible disagreements and implement informed contingency plans (Osuizugbo & Okuntade, 2020).

3. METHODOLOGY

This study adopted a mixed-method research approach to develop an AI-powered predictive model for construction disputes in the Sri Lankan context. A mixed methods design strategically combines quantitative and qualitative research techniques to gain a comprehensive understanding of the complex factors contributing to disputes and to provide robust data for model development. Initial quantitative data collection, using a structured questionnaire and survey, allowed for the identification and measurement of key variables influencing dispute occurrence within historical projects. Qualitative insights, garnered through expert interviews, enriched the understanding of these variables and contextualised the Sri Lankan construction industry's unique dispute landscape. This integrated approach provided a rich dataset, subsequently used to train, optimise, and validate the AI-driven dispute prediction model by having a test run with the practically running projects.

3.1 DATA COLLECTION TECHNIQUES

Expert interviews were employed as a qualitative data collection technique to gain in-depth insights into the complex factors contributing to construction disputes within the Sri Lankan context. These semi-structured interviews targeted industry veterans such as project managers, architects, engineers, and legal professionals, whose extensive experience and specialised knowledge provided valuable perspectives on the causes, consequences, and potential mitigation strategies for construction disputes. The qualitative data derived from these interviews was crucial for enriching the understanding of the quantitative findings gathered through the questionnaire and survey, later strengthening the foundation for the AI model's development.

A meticulously designed questionnaire constituted a primary data collection instrument in this study. The questionnaire was developed based on an extensive literature review and insights gathered from expert interviews, ensuring its focus on critical variables influencing construction disputes in Sri Lanka. A systematic distribution of the questionnaire targeted professionals involved in past or ongoing construction projects, yielding a substantial dataset. This dataset provided quantifiable information regarding project attributes, dispute occurrences, resolution methods, and the perceived impact of various factors on dispute likelihood. The structured nature of the questionnaire allowed for statistical analysis, which played a pivotal role in identifying key patterns and predictors for the AI model.

4. ANALYSIS AND FINDINGS

4.1 METHODS FOR ANALYSING AND DEVELOPING COMPLEX DATA SETS

The set of training data for the machine learning model is generated by applying the underlying logic from the data analysis using Python code. This entails defining construction projects as being the location, type, value, time duration, complexity, and claims of a project. Through extrapolation of these logical rules, the dataset is used in capturing pertinent information that is required in training an AI model for purposes of forecasting the occurrence of potential disputes within construction projects. In line with this, the flow diagram of the proposed model was developed, which outlines the stages of obtaining data on the sources and number of disputes as well as the necessary precautions for utilising clean data and minimising the problem of leakage. The concern with the attributes results from the fact that they are significant in profiling projects and having connections with disputes. The paper has extended the model by adding more parameters that are associated with disputes and these additions further enhance its predictiveness. Organisational dynamics and effectiveness are focal in the development process to provide a dependable information set for future construction disputes. Thus, though there are problems with the efficiency of implemented algorithms, the refined dataset and the training strive to approach errors to manage decisions and risks in constructions.

4.1.1 Findings from Questionnaire Survey

The graph compares the disputed amount with project value, project duration, and complexity (Refer to Figure 1). Analysing this graph led to the following logical findings:

- When the project value increases disputed amount also increases showing a positive relationship.
- Similarly, when the duration of the project increases the disputed amount is also increasing therefore those two attributes show a positive relationship.
- Likewise, when the level of complexity of the project increases, it further affects the increase of the disputed amount showing a positive relationship.
- In conclusion, all the attributes that are considered for drafting the graph have a positive relationship with the disputed amount.

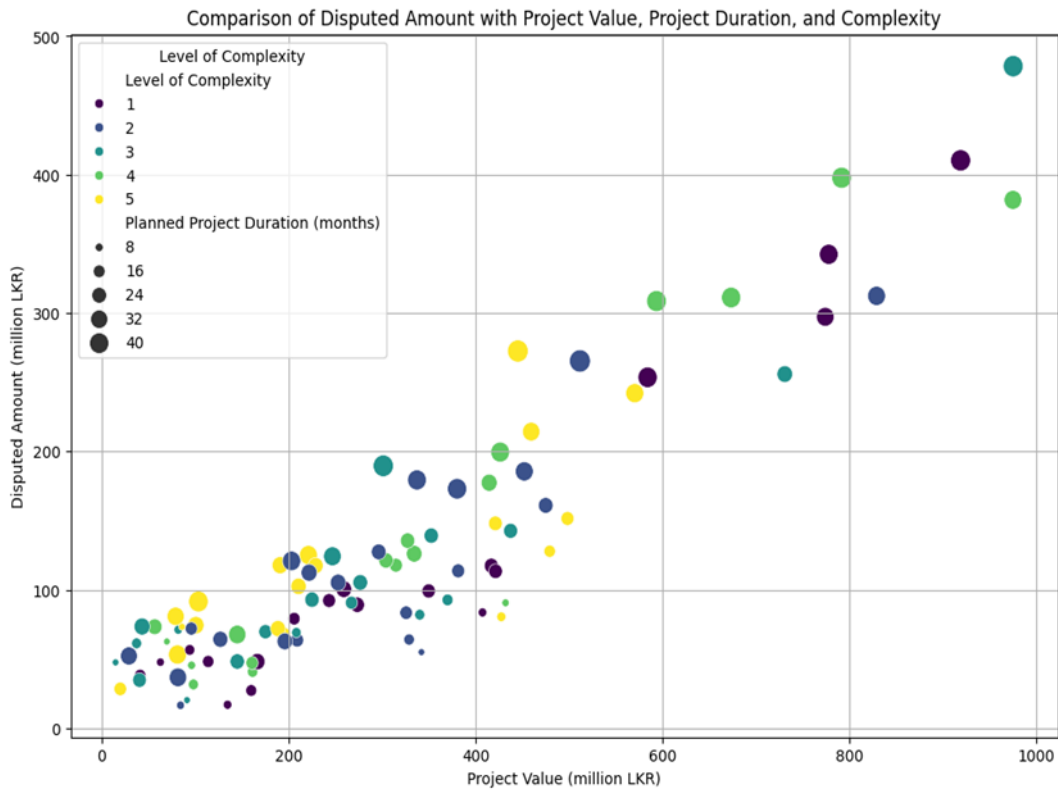


Figure 1: Relationship between disputed amount, project value and complexity

4.1.2 Findings from Expert Interviews

Generic factors that are influenced by the occurrence of construction disputes were analysed through the literature review and expert opinions about the analysed factors were identified using the manual content analysis. These identified factors which are affecting the construction disputes validated by the industry experts, were carried forward to the questionnaire rounds. Then, the finalisation of the conceptual model, intended for use in developing a machine learning model, was undertaken.

4.2 EVOLUTION FROM PRELIMINARY MODEL TO AI MODEL

Figure 2 depicts the sequence of the advancement stages of the model.

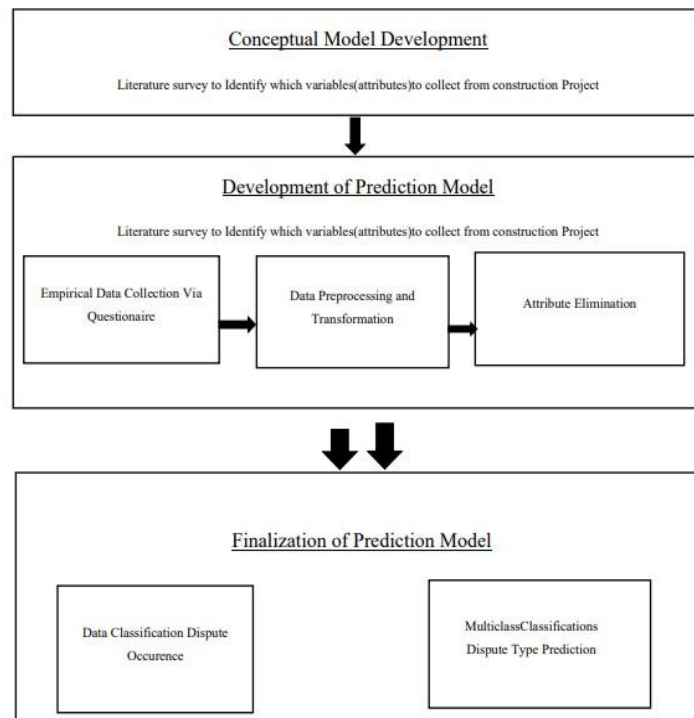


Figure 2: Evolution of models

The stages begin with the ideation phase and gradually progress to the AI implementation stage. This was enabled using well-known data collection techniques such as intensive literature research, and access to a very wide cross-section of experts from within the industry. Each stage of the process involved, was an important milestone that brought about invaluable feedback, many refinements of the model and the model becoming increasingly sophisticated all the time.

This type of iterative journey, as a result, led to the design of a highly satisfying predictive model that represents the intersection of practicality and sophistication. The key point of this approach was the application of data classification machine learning, chosen and adjusted to consider the peculiarities of the domain of a problem. The aim of this algorithmic framework was firstly to maximise the data classification’s potential allowing it to go together with the objective of the AI model. Formulating such a strategy signifies our thorough knowledge coupled with awareness about the complexity existing within the issue space thereby, evoking the need for using the latest computational tools to get access to the operational insight from a data set composed of complex sets. The machine learning framework has been established; this model surpasses the limitations of traditional analytics. As a result, it exhibits the potential for growth and scalability and provides a custom solution for the ever-evolving demands of the construction sector. In summary, it is through a necessary blend of theoretical frameworks, empirical observations, as well as computational wizardry that the models combine state-of-the-art innovation and cutting-edge technology, a testimony to the unreserved dedication to the advancement of predictive analytics.

4.3 FINDINGS FROM LITERATURE REVIEW AND EXPERT INTERVIEW

4.3.1 Features Related to Project Characteristics

Project characteristics and contract-related project characteristics make up the first group of input factors which are affecting for happening construction disputes. The results of the literature review indicate that eleven characteristics fall within this group. Namely, the characteristics are as follows:

- Project location
- Project or contract value
- Purpose of the construction
- Category of the contractor (i.e., joint venture, consortium)
- Type of employer (i.e., public, private, PPP)
- Type of contract
- Method of the payment (i.e., unit price, fixed price)
- Delivery system of the project (i.e., DBB, DB)
- Design complexity of the project
- Construction complexity of the project

4.3.2 Features Related to Skills

The category of the skill includes features based on the stakeholders participating in the project and the characteristics of the organisation. According to the literature survey results, this category includes eight attributes:

- i. Bond between related stakeholders/individuals,
- ii. Years of experience in the construction project (credibility),
- iii. Remedies for avoiding disputes,
- iv. Communication build-up between stakeholders,
- v. Working background cultural features,
- vi. Replying rate and community skills,
- vii. Relevant project experience, and
- viii. Management and coordination skills.

4.3.3 Unexpected Events

The next type of feature selection is 'Changes', which refers to the occurrence of variations, changes, or unforeseen events throughout a construction project. It includes research that discusses the influence of variation on construction conflicts and resolution method selection, as known as the prominence of this variation in the literature, along with the total number of articles that reference them.

4.3.4 Delays

This category examines the influence of delays effect on construction conflicts. It is used to capture the impact of construction project delays. This category considers delays as a factor influencing the type of dispute, and settlement options.

4.3.5 Dispute Background Characteristics

This category of features involves background attributes related to characteristics of a dispute which can be named 'Dispute Characteristics'. Eleven number of attributes were found in this category from the literature survey. These attributes are:

- Dispute affected party
- Dispute happening phase
- Sources that cause the dispute
- Whether there are any suspensions/termination
- Total amount affected by the dispute (financially)
- Agreed amount to resolve the dispute (financially)
- Rate of success (financially)
- Whether there are any EOT claims.
- Amount affected due to EOT
- Agreed settlement for the EOT claim
- Rate of the success of the EOT claim

4.3.6 Method of Resolution Features

This categorisation of features involves characteristics related to the dispute. Therefore, the category is named 'Method of Resolution Features'. There can be numerous characteristics in this category, and they mirror the expectation from the resolution method and order of the selected method. These attributes are:

- Cost for the resolution method selected
- Duration taken by the method selected to resolve the dispute
- Satisfaction level of the resolution method
- Level of importance to preserve relationship
- Speed level of resolving the problem
- Considerable cost regarding solving the problem
- Bindingness importance
- Confidential level of the dispute-resolving method
- Level of fairness in the selected process
- Level of flexibility of the selected process
- Level of control over the selected process
- Level of importance of having the remedy
- Level of importance and willingness of the parties to come into a settlement.

4.3.7 Level of Knowledge Related to Resolution Method

This category mainly focuses on the effect of the potential level of knowledge of the parties who are engaged in the process of decision-making about specific resolution

methods and the selection of the best matching resolution strategy or method in a systematic order.

4.3.8 Machine Learning Model Development Using Conceptual Model

A web application has been developed using the Python framework Flask and a machine-learning model to predict the occurrence characteristics of disputes based on project inputs (refer to Figure 3).

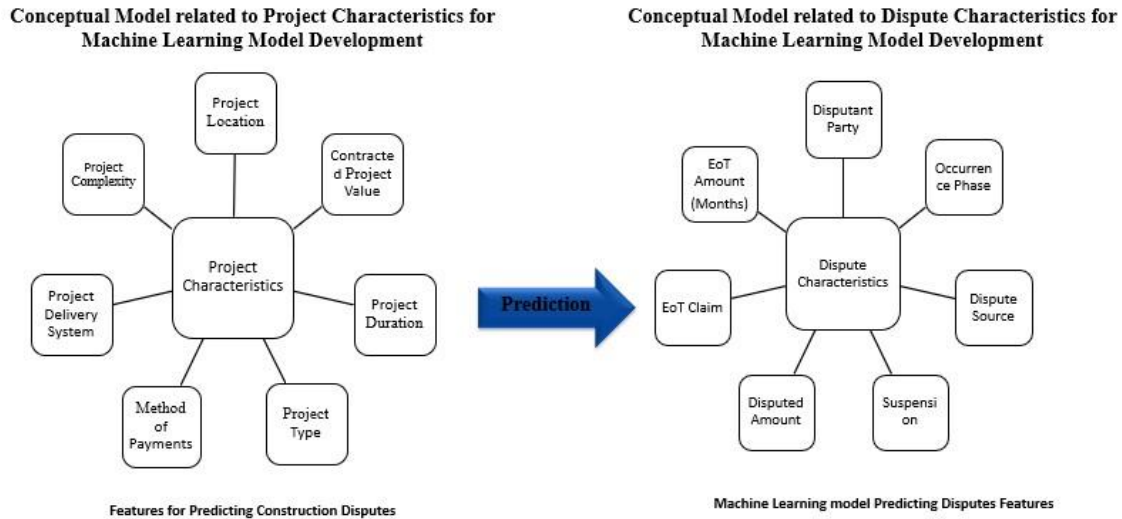


Figure 3: Input data and predicted data

The web application accepts inputs related to project characteristics, which are then passed to the trained machine learning model for prediction. The results from the machine learning model are displayed on the front end of the web application.

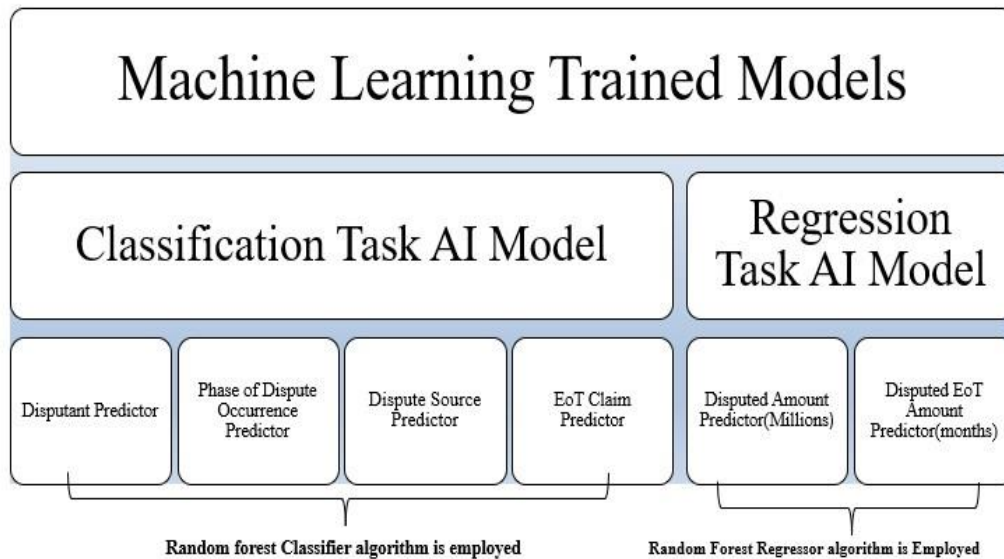


Figure 4: Trained AI models

Training Machine Learning Model

```

# Import necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.metrics import accuracy_score, confusion_matrix
from sklearn.metrics import mean_squared_error, r2_score
import joblib

# Load the dataset
df = pd.read_csv('construction_projects_data.csv')

# Divide categorical variables
label_encoders = {}
for column in df.select_dtypes(include='object').columns:
    le = LabelEncoder()
    df[column] = le.fit_transform(df[column])
label_encoders[column] = le

# Split the data into features and target variables
# For classification: Disputant Party, Phase of Occurrence, Suspended works due to Disputant,
# Disputed Amount, Presence of EoI Claim, Disputed EoI Amount (amount)
y_disputant_party = df['Disputant Party']
y_disputed_amount = df['Disputed Amount']
y_disputed_eoi = df['Disputed EoI Amount']
y_eoi_claim = df['Presence of EoI Claim']
y_disputed_eoi_amount = df['Disputed EoI Amount (amount)']

# For regression: Disputed EoI Amount
# For classification: Disputant Party, Phase of Occurrence, Suspended works due to Disputant,
# Disputed Amount, Presence of EoI Claim, Disputed EoI Amount (amount)
X_train_disputant_party, X_test_disputant_party, y_train_disputant_party, y_test_disputant_party = train_test_split(
    df[['Disputant Party', 'Phase of Occurrence', 'Suspended works due to Disputant', 'Disputed Amount', 'Presence of EoI Claim', 'Disputed EoI Amount (amount)']],
    df['Disputant Party'], test_size=0.2, random_state=42)
X_train_disputed_amount, X_test_disputed_amount, y_train_disputed_amount, y_test_disputed_amount = train_test_split(
    df[['Disputant Party', 'Phase of Occurrence', 'Suspended works due to Disputant', 'Disputed Amount', 'Presence of EoI Claim', 'Disputed EoI Amount (amount)']],
    df['Disputed Amount'], test_size=0.2, random_state=42)
X_train_disputed_eoi, X_test_disputed_eoi, y_train_disputed_eoi, y_test_disputed_eoi = train_test_split(
    df[['Disputant Party', 'Phase of Occurrence', 'Suspended works due to Disputant', 'Disputed Amount', 'Presence of EoI Claim', 'Disputed EoI Amount (amount)']],
    df['Disputed EoI Amount (amount)'], test_size=0.2, random_state=42)
X_train_eoi_claim, X_test_eoi_claim, y_train_eoi_claim, y_test_eoi_claim = train_test_split(
    df[['Disputant Party', 'Phase of Occurrence', 'Suspended works due to Disputant', 'Disputed Amount', 'Presence of EoI Claim', 'Disputed EoI Amount (amount)']],
    df['Presence of EoI Claim'], test_size=0.2, random_state=42)
X_train_disputed_eoi_amount, X_test_disputed_eoi_amount, y_train_disputed_eoi_amount, y_test_disputed_eoi_amount = train_test_split(
    df[['Disputant Party', 'Phase of Occurrence', 'Suspended works due to Disputant', 'Disputed Amount', 'Presence of EoI Claim', 'Disputed EoI Amount (amount)']],
    df['Disputed EoI Amount (amount)'], test_size=0.2, random_state=42)

# Standardize the numerical features for the regression models
scaler = StandardScaler()
X_train_disputed_amount_scaled = scaler.fit_transform(X_train_disputed_amount)
X_test_disputed_amount_scaled = scaler.transform(X_test_disputed_amount)
X_train_disputed_eoi_scaled = scaler.fit_transform(X_train_disputed_eoi)
X_test_disputed_eoi_scaled = scaler.transform(X_test_disputed_eoi)

# Initialize and train the models
# For classification tasks
clf_disputant = RandomForestClassifier(random_state=42)
clf_phase = RandomForestClassifier(random_state=42)
clf_sources = RandomForestClassifier(random_state=42)
clf_suspension = RandomForestClassifier(random_state=42)
clf_eoi_claim = RandomForestClassifier(random_state=42)

clf_disputant.fit(X_train_disputant_party, y_train_disputant_party)
clf_phase.fit(X_train_disputed_amount, y_train_disputed_amount)
clf_sources.fit(X_train_disputed_eoi, y_train_disputed_eoi)
clf_suspension.fit(X_train_disputed_eoi_amount, y_train_disputed_eoi_amount)
clf_eoi_claim.fit(X_train_eoi_claim, y_train_eoi_claim)

# For regression tasks
reg_disputed_amount = RandomForestRegressor(random_state=42)
reg_disputed_eoi = RandomForestRegressor(random_state=42)

reg_disputed_amount.fit(X_train_disputed_amount_scaled, y_train_disputed_amount)
reg_disputed_eoi.fit(X_train_disputed_eoi_scaled, y_train_disputed_eoi)

# Make predictions
y_pred_disputant = clf_disputant.predict(X_test_disputant_party)
y_pred_phase = clf_phase.predict(X_test_disputed_amount_scaled)
y_pred_sources = clf_sources.predict(X_test_disputed_eoi_scaled)
y_pred_suspension = clf_suspension.predict(X_test_disputed_eoi_amount_scaled)
y_pred_eoi_claim = clf_eoi_claim.predict(X_test_eoi_claim)
y_pred_disputed_amount = reg_disputed_amount.predict(X_test_disputed_amount_scaled)
y_pred_disputed_eoi = reg_disputed_eoi.predict(X_test_disputed_eoi_scaled)
    
```

Dividing Dataset in to Training and Testing Dataset

Loading Libraries Pandas, Matplotlib, NumPy, etc....

Classification Type Model

Predicting Inputs & Outputs Variables

Regression Type Model

Predicted Outputs

Figure 5: Training machine learning models

```

# Evaluate the models
acc_disputant = accuracy_score(y_test_disputant, y_pred_disputant)
acc_phase = accuracy_score(y_test_phase, y_pred_phase)
acc_sources = accuracy_score(y_test_sources, y_pred_sources)
acc_suspension = accuracy_score(y_test_suspension, y_pred_suspension)
acc_eoi_claim = accuracy_score(y_test_eoi_claim, y_pred_eoi_claim)
mse_disputed_amount = mean_squared_error(y_test_disputed_amount, y_pred_disputed_amount)
mse_disputed_eoi = mean_squared_error(y_test_disputed_eoi, y_pred_disputed_eoi)

print(f"Accuracy of Disputant Party prediction: {acc_disputant:.2f}")
print(f"Accuracy of Phase of Occurrence prediction: {acc_phase:.2f}")
print(f"Accuracy of Dispute Sources prediction: {acc_sources:.2f}")
print(f"Accuracy of Suspension of Works prediction: {acc_suspension:.2f}")
print(f"Accuracy of Presence of EoI Claim prediction: {acc_eoi_claim:.2f}")
print(f"MSE of Disputed Amount prediction: {mse_disputed_amount:.2f}")
print(f"MSE of Disputed EoI Amount prediction: {mse_disputed_eoi:.2f}")

# Save the models for future use
joblib.dump(clf_disputant, 'clf_disputant.pkl')
joblib.dump(clf_phase, 'clf_phase.pkl')
joblib.dump(clf_sources, 'clf_sources.pkl')
joblib.dump(clf_suspension, 'clf_suspension.pkl')
joblib.dump(clf_eoi_claim, 'clf_eoi_claim.pkl')
joblib.dump(reg_disputed_amount, 'reg_disputed_amount.pkl')
joblib.dump(reg_disputed_eoi, 'reg_disputed_eoi.pkl')

# Save the scalers for future use
joblib.dump(scaler, 'scaler_disputed_amount.pkl')
joblib.dump(scaler, 'scaler_disputed_eoi.pkl')
    
```

Checking the Accuracy of the Trained Model

Saved the Trained Model to Use in the BuildTech Web App

Figure 6: Training machine learning models

4.3.9 Proposed AI Web Application for Dispute Prediction

The recommended solution for applying AI to predict construction disputes is a construction dispute management system designed to serve as an intermediary for dispute avoidance. The system should include:

- Login page where the user can log in to the system using his/her email and a password.
- Registration page where a user can register for the system and can have access to other features in the dispute management system.
- Dashboard which showcases the performance or the accuracy of the dispute prediction model.
- Interface consisted of the reports which are prepared for the project.

- Interface which gives brief details about the project in which the system login party is involved.
- Dispute prediction page where the disputed amount, dispute source, and the dispute occurrence phase are predicted using other variables.
- ‘Build Tech’ is the suggested name for the dispute management system.

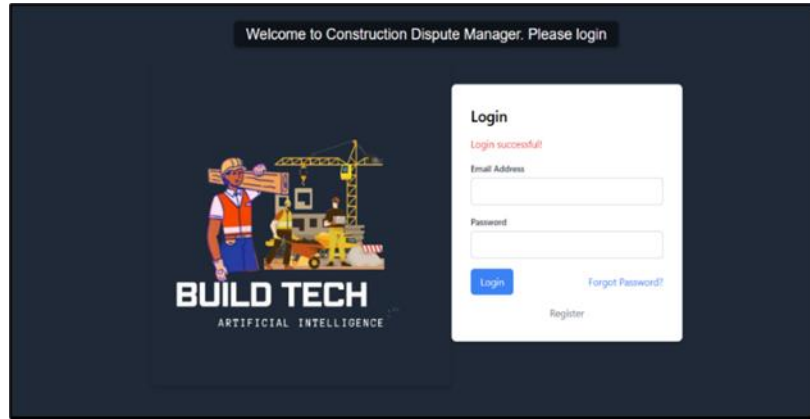


Figure 7: Proposed Flask web app solution

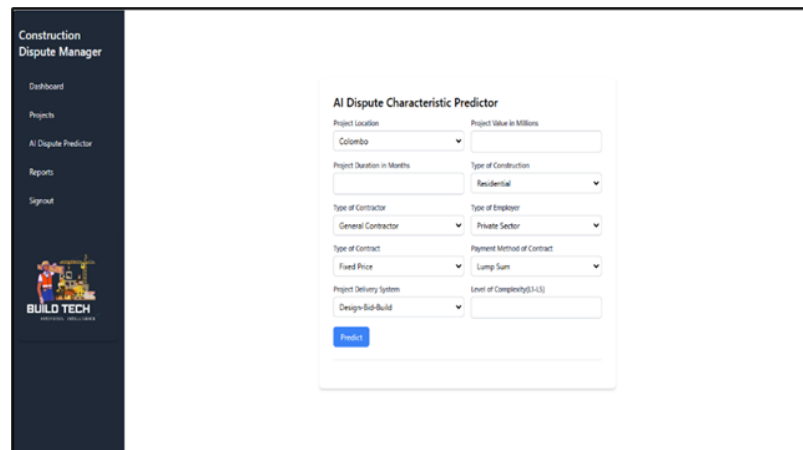


Figure 8: Proposed Flask web app solution

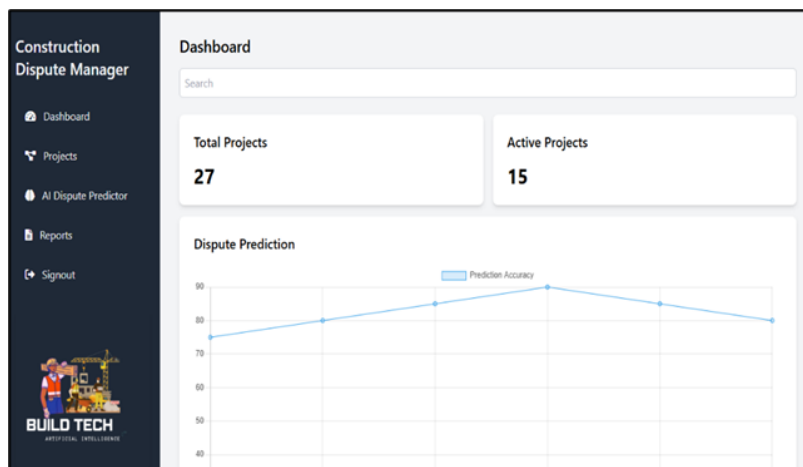


Figure 9: Proposed Flask web app solution

5. CONCLUSIONS

The study has facilitated the successful research of an innovative predictive model for construction disputes in Sri Lanka, which is incorporated with AI technology in the first life cycle phase of a project. The study adopted a mixed-measurement research design where the author conducted a cross-sectional survey and interviewed technical experts in the Sri Lankan construction industry to explore potential sources of dispute.

The predictive model has been calibrated considering the potential of adversarial situations given features that are associated with conflict risks such as breached contractual terms, unrealistic schedules, payment issues, communication breakdowns, and adverse ground conditions. It can aid construction project stakeholders to adopt effective and timely strategies to avoid or minimise possible disputes and enhance the chances of optimising opportunities of post-improvement for the general betterment of the projects, in terms of possible time and cost overrun, as well as deterioration of friendly professional relations.

The long-term implications of this research are significant, as the widespread adoption of this AI model has the potential to reshape dispute management practices and promote a more collaborative and sustainable construction industry in Sri Lanka. By communicating early indicators of impending disputes to project teams, this advanced tool provides an opportunity for proactive response and the implementation of appropriate measures to prevent conflict escalation. This contributes to the successful completion of construction projects in the country.

6. REFERENCES

- Abeyasinghe, N., & Jayathilaka, R. (2022). Factors influencing the timely completion of construction projects in Sri Lanka. *PLOS ONE*, *17*. <https://doi.org/10.1371/journal.pone.0278318>
- Agus, A., Sudirman, S., Umar, W., & Rustan, A. (2023). The use of artificial intelligence in dispute resolution through arbitration: The potential and challenges. *SASI*, *29*(3), 570. <https://doi.org/10.47268/sasi.v29i3.1393>
- Arfazadeh, H. (2014). Arbitrability under the New York convention: the Lex Fori revisited, *Arbitration International*, *17*(1) 73–88. <https://doi.org/10.1023/A:1008994201415>
- Assaf, S., Hassanain, M. A., Abdallah, A., Sayed, A. M. Z., & Alshahrani, A. (2019). Significant causes of claims and disputes in construction projects in Saudi Arabia. *Built Environment Project and Asset Management*, *9*(5), 597–615. <https://doi.org/10.1108/BEPAM-09-2018-0113>
- Carcillo, F., Le Borgne, Y. A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, *557*, 317–331. <https://doi.org/10.1016/j.ins.2019.05.042>
- Caruana, R. A. (1993, June 27). *Multitask learning: A knowledge-based source of inductive bias*. *Machine Learning*, *28*, 4175. <https://doi.org/10.1016/b978-1-55860-307-3.50012-5>
- Choi, S. J., Choi, S. W., Kim, J. H., & Lee, E. B. (2021). Ai and text-mining applications for analyzing contractor's risk in invitation to bid (ITB) and contracts for engineering procurement and construction (EPC) projects. *Energies*, *14*(15), 4632. <https://doi.org/10.3390/en14154632>
- Chou, J. S., Cheng, M. Y., Wu, Y. W., & Pham, A. D. (2014). Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification. *Expert Systems with Applications*, *41*(8), 3955–3964. <https://doi.org/10.1016/j.eswa.2013.12.035>
- Collier, P. (2020). Embedding the sustainability development goals in Survey Review's remit. *Survey Review*, *52*(373), 287-288. <https://doi.org/10.1080/00396265.2020.1746020>
- El Naqa, I., & Murphy, M. J. (2015). What is machine learning? In *Machine learning in radiation oncology*. Springer International Publishing. https://doi.org/10.1007/978-3-319-18305-3_1

- El-Sayegh, S., Ahmad, I., Aljanabi, M., Herzallah, R., Metry, S., & El-Ashwal, O. (2020). Construction disputes in the UAE: Causes and resolution methods. *Buildings*, *10*(10), 171. <https://doi.org/10.3390/buildings10100171>
- Fei, W., Opoku, A., Agyekum, K., Oppon, J. A., Ahmed, V., Chen, C., & Lok, K. L. (2021). The critical role of the construction industry in achieving the sustainable development goals (Sdgs): Delivering projects for the common good. *Sustainability*, *13*(16). <https://doi.org/10.3390/su13169112>
- Gunarathna, C., Yang, R. J., & Fernando, N. G. (2018). Conflicts and management styles in the Sri Lankan commercial building sector. *Engineering, Construction and Architectural Management*, *25*(2), 178-201. <https://doi.org/10.1108/ECAM-10-2016-0233>
- Horvitz, E., & Mulligan, D. (2015). Data, privacy, and the greater good. *Science*, *349*(6245), 253–255. <https://doi.org/10.1126/science.aac4520>
- Ikechukwu, A. C., Emoh, F. I., & Kelvin, O. A. (2017). Causes and effects of cost overruns in public building construction projects delivery, in Imo State, Nigeria. *Journal of Business and Management*, *19*(7), 13–20. <https://doi.org/10.9790/487x-1907021320>
- Liu, B. (2017). Lifelong machine learning: A paradigm for continuous learning. *Frontiers of Computer Science*, *11*(3), 359–361. <https://doi.org/10.1007/s11704-016-6903-6>
- Marciano, A., & Ramello, G. B. (2019). Introduction to the symposium on the empirics of judicial institutions. *Journal of Institutional Economics*, *15*(1), 73–80. <https://doi.org/10.1017/S1744137418000322>
- Mashwama, N., Thwala, W. D., & Aigbavboa, C. O. (2019). The impact of construction disputes on projects in the Mpumalanga province of South Africa. In J.S. Miroslaw & M. Hajdu (Eds.), *Proceedings of the Creative Construction Conference* (pp. 454–461). <https://doi.org/10.3311/CCC2019-063>
- Molu, M. M., & Goertz, N. (2014). A comparison of soft-coded and hard-coded relaying. *Transactions on Emerging Telecommunications Technologies*, *25*(3), 308–319. <https://doi.org/10.1002/ett.2562>
- Osuizugbo I. C., & Okuntade, T. F. (2020). Conflict management practice among stakeholders in construction project delivery. *Covenant Journal in Research & Built Environment*, *8*(1). <http://journals.covenantuniversity.edu.ng/index.php/cjrbe>
- Sandagomika, N. M. G. H., Sandanayake, Y., & Ekanayake, Biyanka. (2020). Drivers and barriers of using Internet of Things for successful lean implementation in construction projects in Sri Lanka. *FARU Journal*, *7*(12). <https://doi.org/10.4038/faruj.v7i0.26>
- Serogina, D., Pushkar, T., & Zhovtiak, H. (2022). The impact of the construction industry on the social and economic development of territories. *Scientific Bulletin of the National Academy of Statistics, Accounting and Audit*, *1*(2), 59–65. <https://doi.org/10.31767/nasoa.1-2-2021.08>
- Silva, G. A. S. K., Warnakulasuriya, B. N. F., & Arachchige, B. J. H. (2018). A review of the skill shortage challenge in construction industry in Sri Lanka. *International Journal of Economics, Business and Management Research*, *2*(1). https://www.researchgate.net/publication/320263862_a_review_of_the_skill_shortage_challenge_in_construction_industry_in_sri_lanka
- Walsh, T. W., & Llp, C. (2007). The LCIA court decisions on challenges to arbitrators: An introduction. *Arbitration International*, *1*(3). <http://arbitration.oxfordjournals.org/>
- Wang, L., Zhang, H., Wang, J., & Li, L. (2018). Picture fuzzy normalized projection-based VIKOR method for the risk evaluation of construction project. *Applied Soft Computing Journal*, *64*, 216–226. <https://doi.org/10.1016/j.asoc.2017.12.014>
- Yildizel, S. A., Dogan, E., Kaplan, G., & Ergut, A. (2016). Major constructional dispute causes in Turkey. *Archives of Civil Engineering*, *4*(2). <http://archive.sciendo.com/ACE/ace.2016.62.issue-4/ace-2015-0116/ace-2015-0116.pdf>
- Zhou, Z. H. (2018). A brief introduction to weakly supervised learning. In *National Science Review*. Oxford University Press. <https://doi.org/10.1093/nsr/nwx106>