MCSD 1043
RESEARCH DESIGN AND ANALYSIS IN DATA SCIENCE

PROJECT PROPOSAL PRESENTATION

TEMPORAL ANALYSIS OF CLIMATIC INFLUENCES ON FOREST FIRE PATTERNS IN PENINSULAR MALAYSIA USING STATISTICAL METHOD

CANDIDATE : GRACE LING KIAN HWAI
LECTURER : ASSOC. PROF. DR MOHD SHAHIZAN BIN OTHMAN

VENUE : DISCUSSION ROOM 4, LEVEL 2, BLOCK N28A
DATE : 25 JULY 2024

FACULTY OF COMPUTING,
UNIVERSITI TEKNOLOGI MALAYSIA

www.utm.my
innovative • entrepreneurial • global

univteknologimalaysia utm_my utmofficial

PRESENTATION CONTENTS

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

TEMPORAL ANALYSIS OF CLIMATIC INFLUENCES ON FOREST FIRE PATTERNS IN PENINSULAR MALAYSIA USING STATISTICAL METHOD

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

RESEARCH INTRODUCTION

INTRODUCTION

PROBLEM BACKGROUND AND STATEMENT

Problem Background

Previous Studies (Globally)

- Increased temperature and decreased precipitation contribute to fire frequency and severity. Turco et al. (2014)
- Historical data analysis showing the impact of temperature and precipitation changes on fire frequency and severity. Flannigan et al. (2005)
- Dynamic vegetation modeling in Alaskan boreal habitats predicting increased fire size and frequency under various climate conditions. Balshi et al. (2009)

Problem Statement

Previous Studies (Malaysia)

- A thorough review and analysis has not yet been undertaken in Malaysia. (Chew et al., 2022)

INTRODUCTION

THE NEED OF STUDY

961 forest and bush fires Sabah

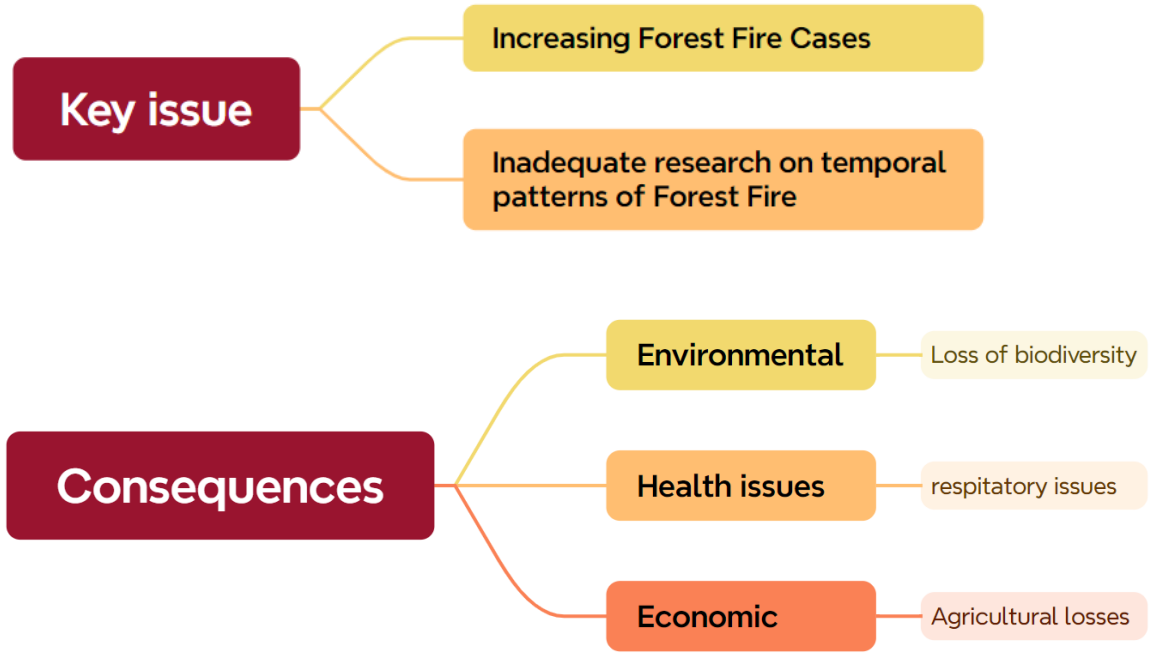
Thursday, 04 Apr 2024 8:22 PM MYT

KOTA BARU, April 4 — The Kelantan Fire and Rescue Department recorded 577 fire cases in the first three months of this year, compared to 257 cases during the same period last year.

By Ben Tan

Tuesday, 17 Oct 2023 10:36 PM MYT

JOHOR BARU, Oct 17 — The Johor Fire and Rescue Department has successfully extinguished a fire involving 23 hectares of forest in Pengerang near Kota Tinggi after three days of operation.



Need for Study!

Urgent requirement for comprehensive data and analysis to improve forest fire management and mitigation strategies.

INTRODUCTION

RESEARCH QUESTIONS

Research Questions

1. What are the patterns and trends can be deduced in relation to the forest fire incidences in Peninsular Malaysia in the last five years?

2. Which month and seasons of the year and where are the hotspot areas responsive to the highest number of times that forest fire occurred in Peninsular Malaysia?

3. What is the relationship between geographical characteristic in the Peninsular Malaysia and the climatic factors that determine forest fire?

INTRODUCTION

RESEARCH OBJECTIVES

Research Objectives

1. To study the forest fire patterns and climatic features that coincide with the period of the analysis in Peninsular Malaysia.
2. To analyse the temporal characteristics of the forest fire occurrences and determine the relationship between climatic variables and forest fire occurrences.
3. To create models that will allow to predict the factors of forest fire potential depending on climate.

INTRODUCTION

RESEARCH MAPPING

	Problem Statements	Research Questions	Research Objectives
1	Lack of understanding of recent trends in forest fire incidences, hindering effective forest fire management strategies.	What are the patterns and trends can be deduced in relation to the forest fire incidences in Peninsular Malaysia in the last five years?	To study the forest fire patterns and climatic features that coincide with the period of the analysis in Peninsular Malaysia.
2	Uncertainty about the temporal patterns of forest fire occurrences, making it difficult to allocate resources and implement preventive measures effectively.	Which month and seasons of the year and where are the hotspot areas responsive to the highest number of times that forest fire occurred in Peninsular Malaysia?	To analyse the temporal characteristics of the forest fire occurrences and determine the relationship between climatic variables and forest fire occurrences.
3	Lack of detailed knowledge about how geographical and climatic factors interact to influence forest fire occurrences, impeding the development of targeted mitigation strategies.	What is the relationship between geographical characteristics in Peninsular Malaysia and the climatic factors that determine forest fire?	To create models that will allow predicting the factors of forest fire potential depending on climate.

INTRODUCTION

RESEARCH SCOPE

Scope

1. The study will utilize the dataset provided by the collaborator where the **dataset** is obtained from **Global Remote Sensing Data from Google Earth Engine (GEE).**
2. The analysis will cover **at least 5 years** to capture long-term trends and patterns
3. Key climatic variables such as **temperature, humidity and windspeed** will be examined for their influence on forest fire patterns.
4. The study will focus exclusively on **Peninsular Malaysia**, excluding East Malaysia due to different climatic and ecological conditions.
5. The research will use **historical data that is currently obtainable**, although it may have some constraints regarding precision and comprehensiveness.
6. **Statistical and machine learning tools** will be used for data analysis and model development.

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

LITERATURE REVIEW

LITERATURE REVIEW

KEY COMPONENTS

Key Components

Statistical Methods

Definition	Mathematical techniques for data analysis
Applications	- Healthcare - Economics - Environmental studies
Research Evolution	**- Early techniques:** Regression analysis, hypothesis testing (Cameron & Trivedi, n.d.) **- Recent advancements:** Machine learning algorithms, Bayesian statistics (Gelman et al., n.d.)
Techniques	- Descriptive and inferential statistics, parametric and non-parametric tests (Luz et al., n.d.). - Linear regression, ANOVA (James et al., 2021)

Temporal Analysis

Definition	Study of data pattern over time
Applications	- Climatology - Ecomnomics - Social sciences
Advancements	**- Early models:** ARIMA, exponential smoothing (G. Jain, n.d.) **- Recent models:** LSTM networks, Temporal Convolutional Networks (TCNs) (Gopali et al., 2021)
Techniques	- Time series analysis (Belay et al., 2021) - Spatial-temporal analysis (Liu & Song, 2020) - Temporal network analysis (Cencetti et al., 2021)

Peninsular Malaysia

Geographical Context	West Malaysia, located between Thailand, Strait of Malacca, South China Sea and Singapore (Olaniyi et al., 2011)
Applications	- Deforestation studies - Socio-economic effects of tourism - Climate change adaptation
Research Evolution	**- Early focus:** Colonial history (Ngah, n.d.) **- Recent focus:** Sustainable development, environmental conservation, socio-economic challenges (Laporan Akhir Penyelidik, n.d.)
Methodologies	- Field surveys (Khalid, 2018) - Remote sensing (Leman et al., 2016) - GIS mapping (Leman et al., 2016). - Statistical analysis (Khalid, 2018)

Forest Fire Patterns

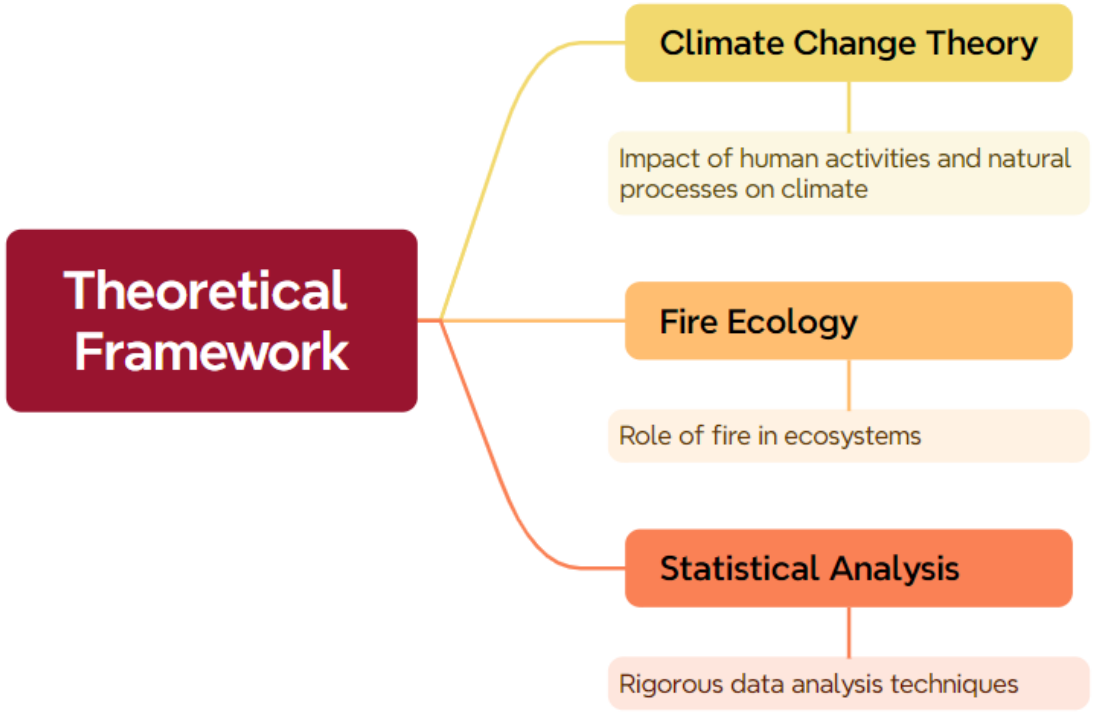
Definition	Spatial and temporal occurrence of wildfires
Applications	- Fire management - Policy-making
Research Evolution	**- Early studies:** Ground-based observations (Mangeon et al., 2016) **- Recent advancements:** Remote sensing, GIS, (Szpakowski & Jensen, 2019), machine learning(P. Jain et al., 2020)
Techniques	- Remote sensing (Giglio et al., n.d.) - Climate modeling (Abatzoglou & Williams, 2016) - ML (P. Jain et al., 2020)

Climate Influences

Definition	Impact of climate conditions on natural and human systems
Applications	- Agriculture - public health - environmental conservation
Research Evolution	**- Early studies:** observational studies (Treut et al., n.d.) **- Recent advancement:** advanced climate models and satellite technology (Fatichi et al., 2011)
Methodologies	- Statistical analysis - Climate modeling (Robock et al., 1993) - Remote sensing (Thies & Bendix, 2011) - ML (Ford et al., 2016)

LITERATURE REVIEW

THEORETICAL FRAMEWORK



Study	Statistical Method	Climatic Variables	Fire Pattern Identified
Dastour et al. (2024)	Z-score Analysis	Trend component of RobustSTL	Identification of anomalous periods in time series data
	Mann-Kendall (MK) Test	Relative Humidity, Precipitation, Air Temperature (Min & Max)	Detection of trends in climatic variables and fire occurrences
	Sen's Slope Estimator (SSE)	Relative Humidity, Precipitation, Air Temperature (Min & Max)	Evaluation of the magnitude and direction of trends
Halofsky et al. (2020)	Statistical Model	Temperature, Precipitation, Global Climate Models	Future projections of area burned; Increased area burned with warming climate
P. Jain et al. (2020)	Self-Organizing Maps (SOMs)	Temperature, Humidity, Rainfall	Correlation between fire weather and large-scale climatic patterns
	Linear Regression (LR)	Temperature, Rainfall	Predicting fire danger indices based on weather observations
	Random Forest (RF)	Various climatic variables	Superior performance in fire severity mapping and fire detection
Szpakowski & Jensen (2019)	Point-wise meteorological data-based operating systems (WFAS, FWI, FFDRS, Nesterov Index)	- Temperature - Humidity - Precipitation - Wind speed	- Fire hazard mapping based on environmental factors like fuel conditions and topography - Dynamic variables like fuel moisture and vegetation conditions for short-term fire risk mapping
	Remote sensing and GIS techniques		

LITERATURE REVIEW

RESEARCH GAP

Aspect	Previous Studies	Current Research	Research Gap
Data Sources	Satellite imagery, weather stations, historical records	Combination of satellite imagery, climate models, and field data	Integration of multiple data sources for comprehensive analysis
Methodology	Regression analysis, predictive modelling, remote sensing	Advanced statistical methods	Application of machine learning for improved prediction accuracy
Focus Area	General climatic influences on fires	Specific focus on Peninsular Malaysia	Lack of region-specific studies addressing local climatic conditions
Policy Recommendations	General fire management strategies	Tailored recommendations for Peninsular Malaysia	Need for localized strategies based on detailed analysis

Giglio et al. (2020)

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

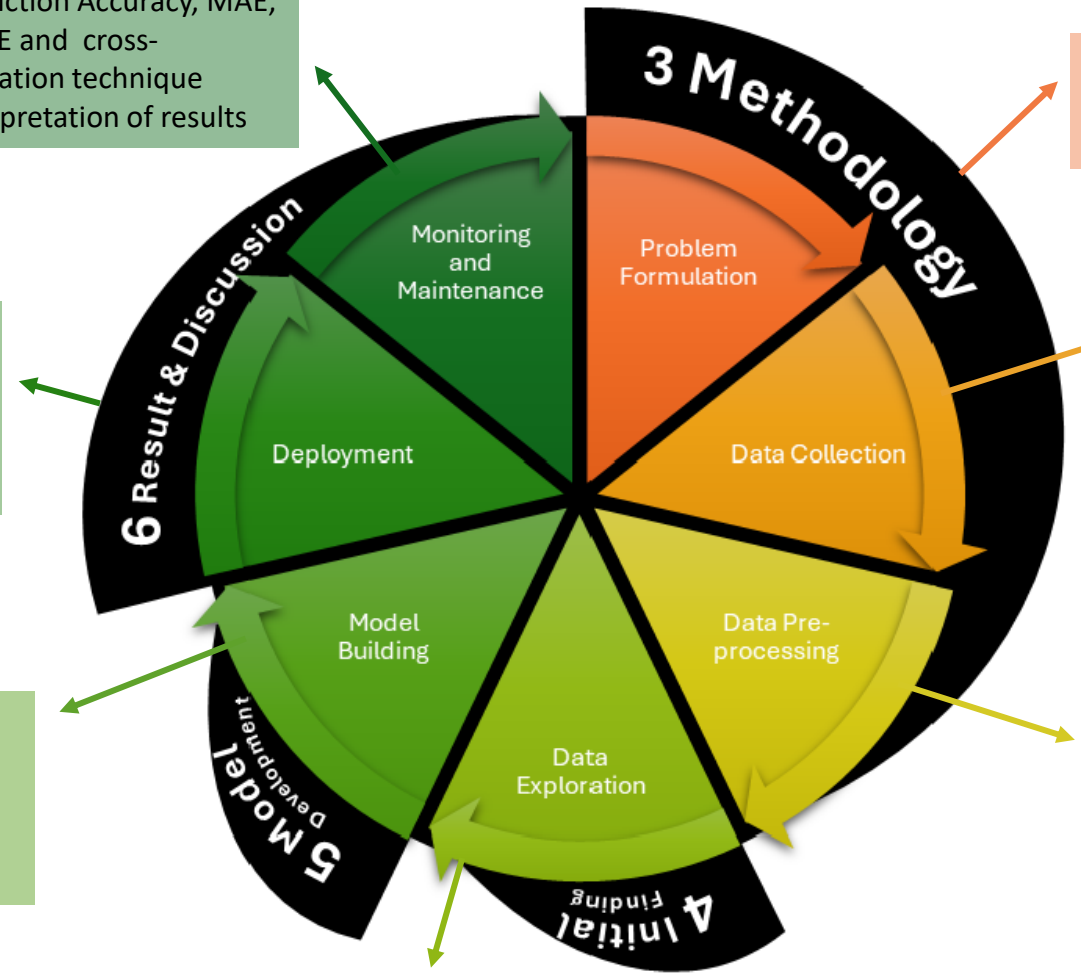
RESEARCH METHODOLOGY

RESEARCH METHODOLOGY

RESEARCH FRAMEWORK

Monitoring & Maintenance:
Performance Evaluation

- Evaluate the performance of each model by Prediction Accuracy, MAE, RMSE and cross-validation technique
- Interpretation of results



Problem Formulation
Problem Definition & Literature Review

Data Collection
Obtained forest fire and climate change data from 2001-2023 from Collaborator which extracted from Multiple-Source Remote Sensing Data using Google Earth Engine.

Data Pre-processing

1. **Preliminary Analysis**
   - Loading the Data
   - Initial Exploration
   - Checking for Missing and Duplicate Values
2. **Data Wrangling**
   - Concatenating Monthly Data
   - Date Conversion
   - Filtering Data
   - Handling Missing and Duplicate Data
   - Data Transformation

Deployment :
Testing and Validation

- Test both models based on testing set
- Tuning Hyperparameters

Model Building

- Feature Engineering
- Develop Statistical Model (Regression Model, Time Series Analysis)

Data Exploration
Perform Exploratory Data Analysis

- Visualization
- Statistical Analysis

RESEARCH METHODOLOGY

DATASET

Primary Data

Dataset size : 82.89M
Dataset covered period : 2001 - 2023

Forest Fire Dataset

Source: Remote sensing data via Google Earth Engine

Details: Location, time, intensity of fires

Feature Name	Description
system:index	System-generated from MCD64A1
longitude	Longitude Coordinate of Fire Points
latitude	Latitude Coordinate of Fire Points
fire	Fire Occurrence (binary class)
date	Date from Administrative Boundaries refer to the Shape
ADM1_PCODE	Administrative level 1 code
ADM2_PCODE	Administrative level 2 code
Shape_Leng	Shape Length (from MCD64A1)
ADM0_EN	Country Name
ADM1_EN	Administrative level 1 name
ADM2_EN	Administrative level 2 name
validOn	Validation Date from Administrative Boundaries refer to the Shape
Shape_Area	Shape area (from MCD64A1)
ADM0_PCODE	Country code
BurnDate	Date in 0–365 (from MCD64A1)
year	Year of Fire Observation
month	Month of Fire Observation
day	Day of Fire Observation
current0101_hii_annual	Human Impact Index
current0101_average_annual_nighttime	Nighttime Brightness

Climate Change Dataset

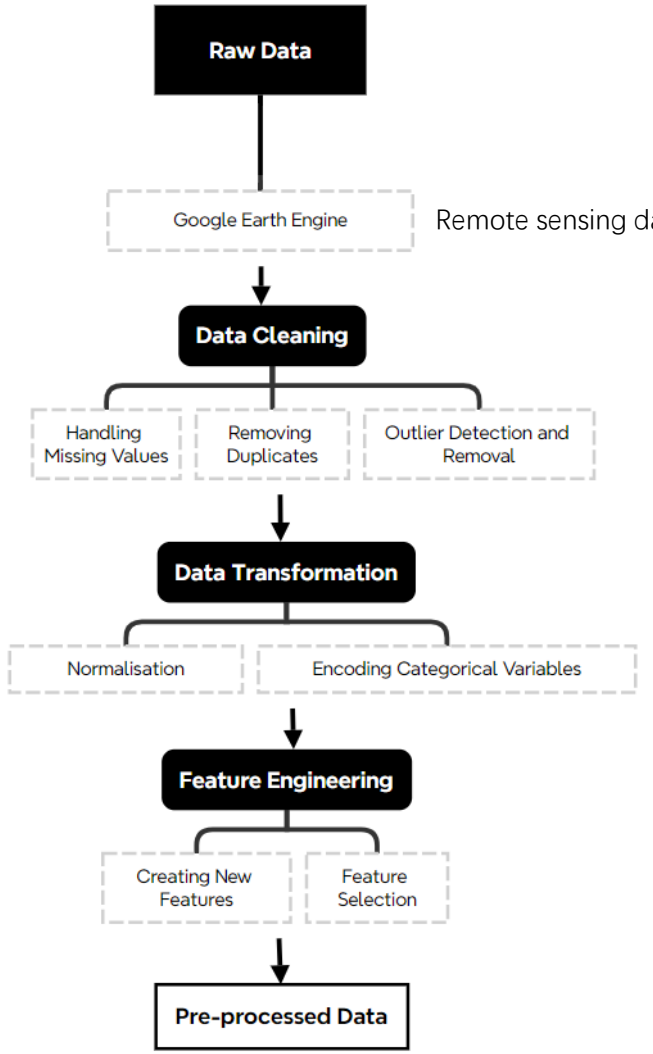
Source: Remote sensing data via Google Earth Engine

Variables: Temperature, moisture, wind velocity

Feature Name	Description
current_aet_annual	Actual Evapotranspiration
current_def_annual	Climate Water Deficit
current_pdsi_annual	Palmer Drought Severity Index
current_pet_annual	Reference Evapotranspiration
current_pr_annual	Precipitation Accumulation
current_ro_annual	Runoff
current_soil_annual	Soil Moisture
current_srad_annual	Downward Surface Shortwave Radiation
current_swe_annual	Snow Water Equivalent
current_tmmn_annual	Minimum Temperature
current_tmmx_annual	Maximum Temperature
current_vap_annual	Vapor Pressure
current_vpd_annual	Vapor Pressure Deficit
current_vs_annual	Wind Speed at 10 m
current_EVI_annual	Enhanced Vegetation Index
current_NDVI_annual	Normalized Difference Vegetation Index
current_LST_annual	Land Surface Temperature
current_KBDI_annual	Keetch–Byram Drought Index
current0101_LC_Type2_annual	Land Cover Classification of UMD (Numeric)
current0101_LC_Type2_annual_classname	Land Cover Classification of UMD (Class Name)

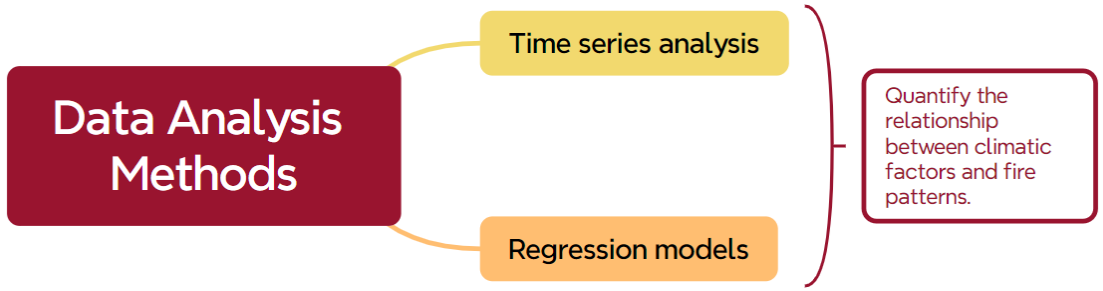
RESEARCH METHODOLOGY

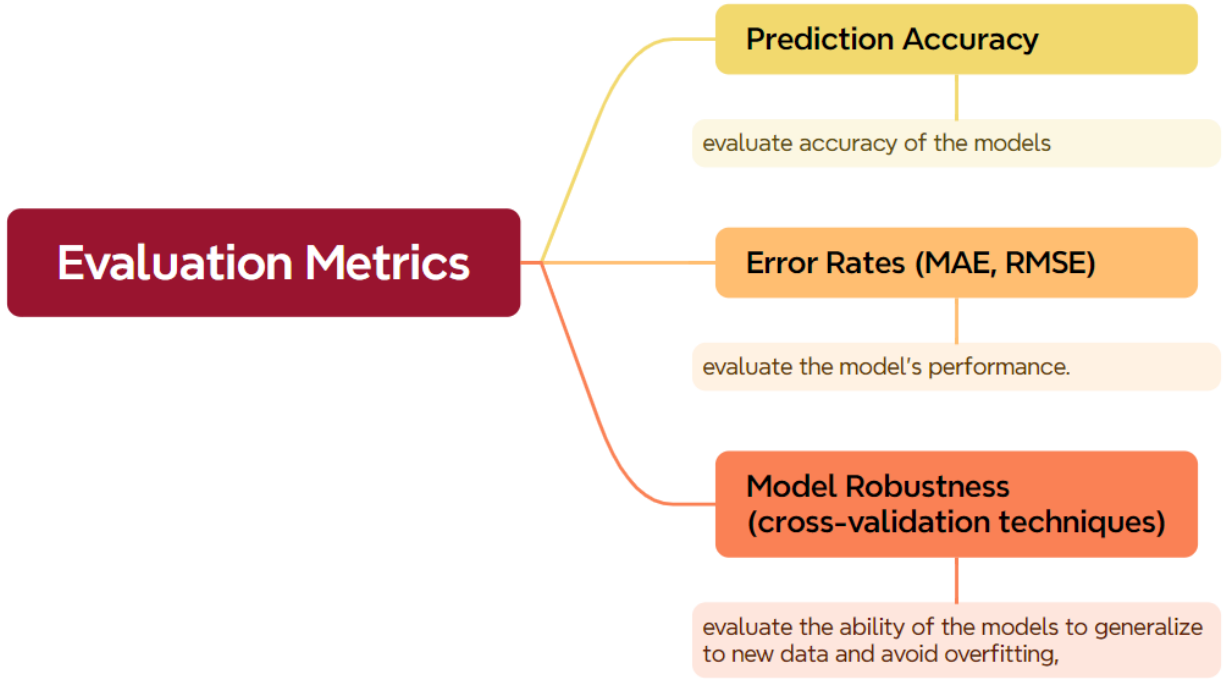
DATA PREPROCESSING



RESEARCH METHODOLOGY

DATA ANALYSIS AND EVALUATION METRICS





RESEARCH METHODOLOGY

COMPARISON OF METHODS

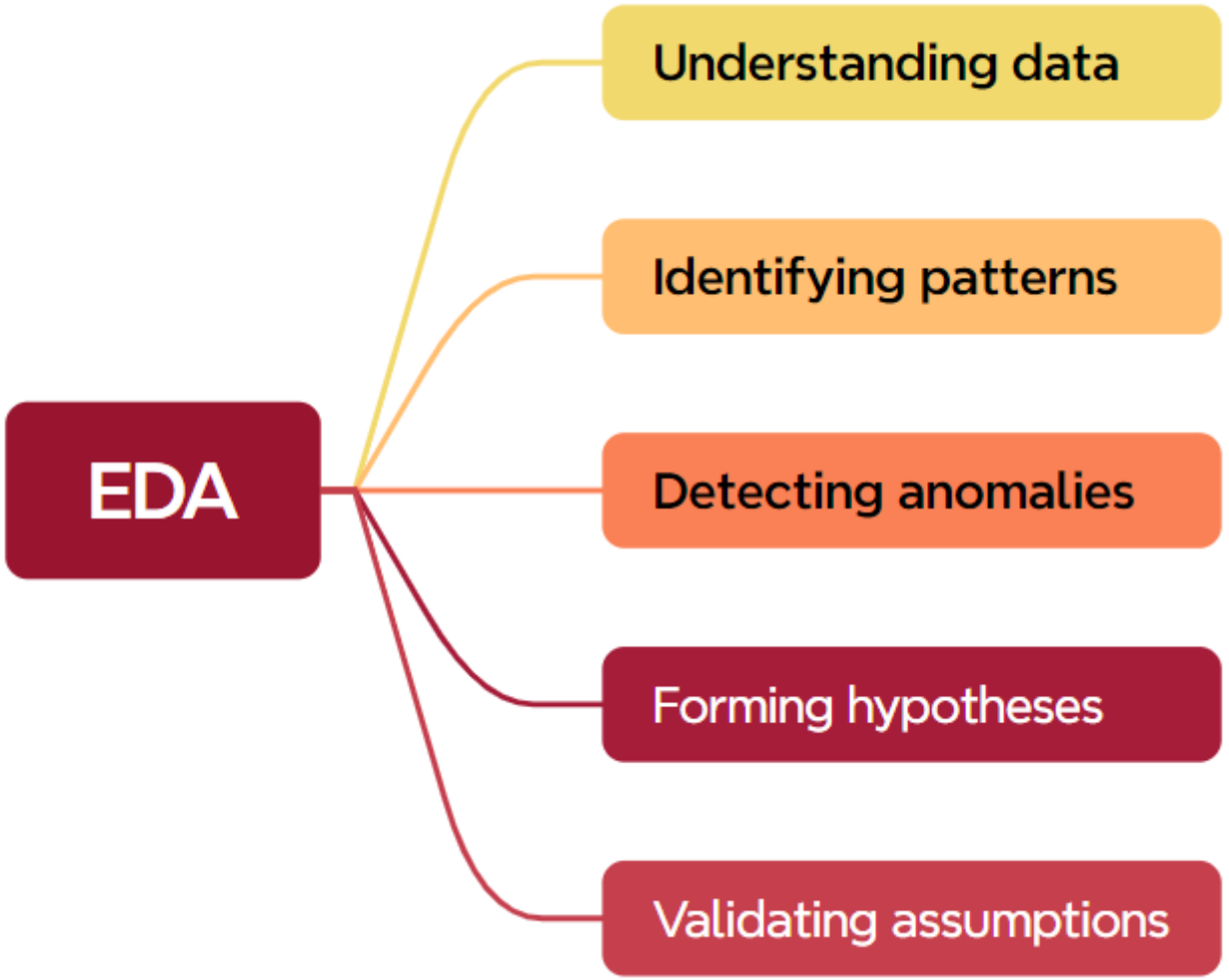
Aspect	Previous Studies	Current Research	Research Gap
Data Collection	Satellite imagery, weather stations	Google Earth Engine, IMF Climate Change Dashboard	Need for high-resolution temporal data
Methodology	Statistical analysis, machine learning	Advanced statistical methods	Integration of advanced AI techniques
Geographic Focus	Southeast Asia	Peninsular Malaysia	Need for region-specific studies
Validation	Separate dataset	Cross-validation with real-time data	Implementation of real-time validation techniques
Applications	Fire detection, prediction	Fire detection, prediction, management	Development of comprehensive fire management systems

Sabani et al. (2009);
Saruni Dwiasnati & Yudo Devianto (2021);
Ghali and Akhloufi (2023);
Sudiana et al. (2023);
Meng et al. (2024)

INITIAL FINDINGS

INITIAL FINDINGS

Exploratory Data Analysis (EDA)

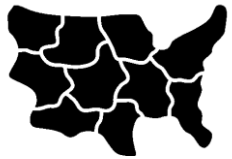


INITIAL FINDINGS

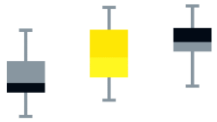
VISUALIZATIONS

Time Series Plots: Displaying the temporal patterns of forest fire occurrences and climatic variables.

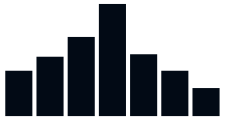
Heatmaps: Showing the correlation between different climatic variables and fire incidents.



Geographical Maps: Highlighting the spatial distribution of forest fires across Peninsular Malaysia.



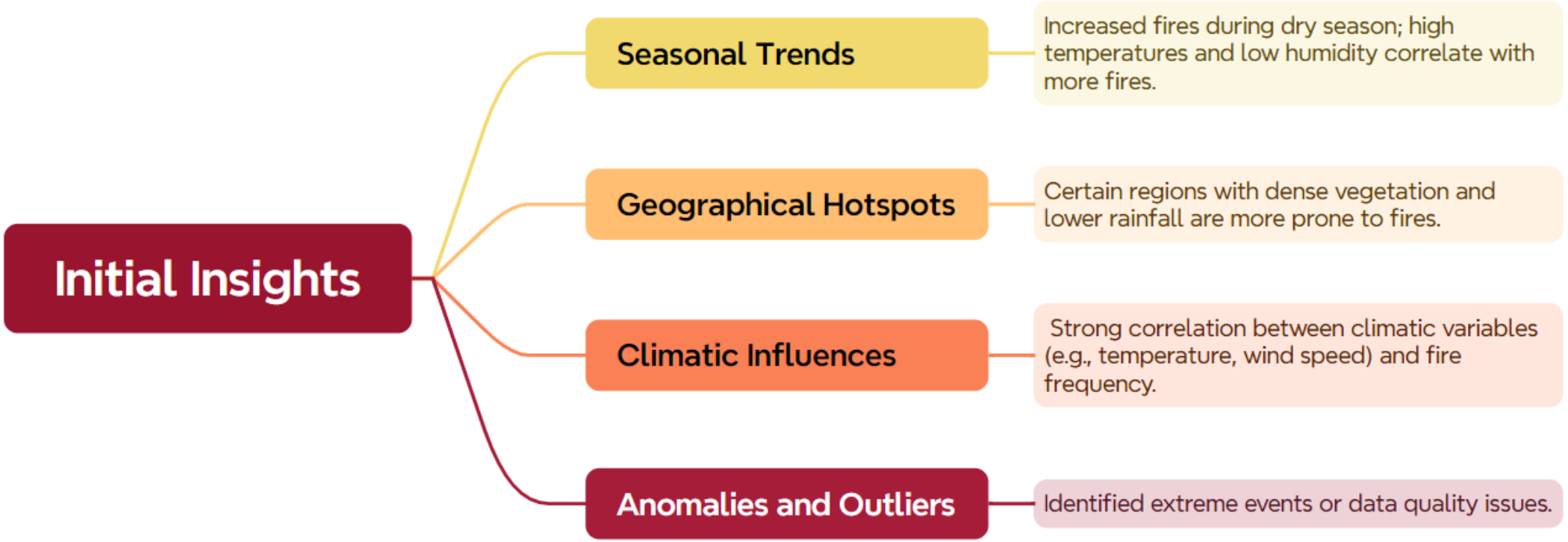
Box Plots: Illustrating the distribution and range of climatic variables during fire and non-fire periods.



Histograms: Depicting the frequency distribution of fire occurrences and climatic variables.

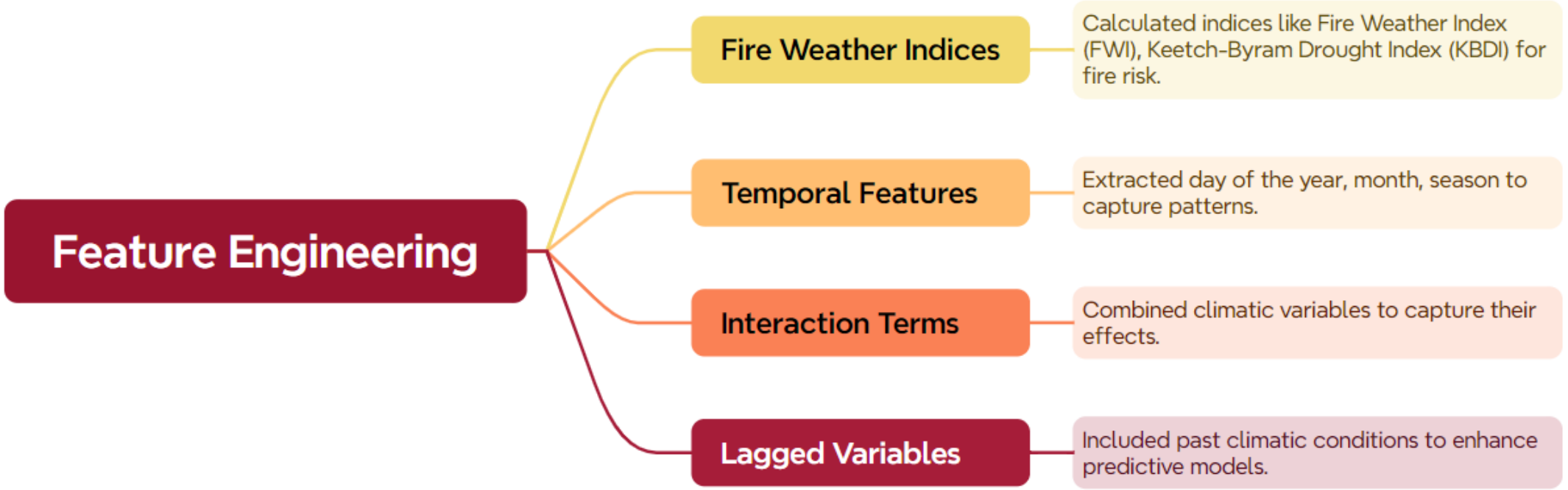
INITIAL FINDINGS

INITIAL INSIGHTS



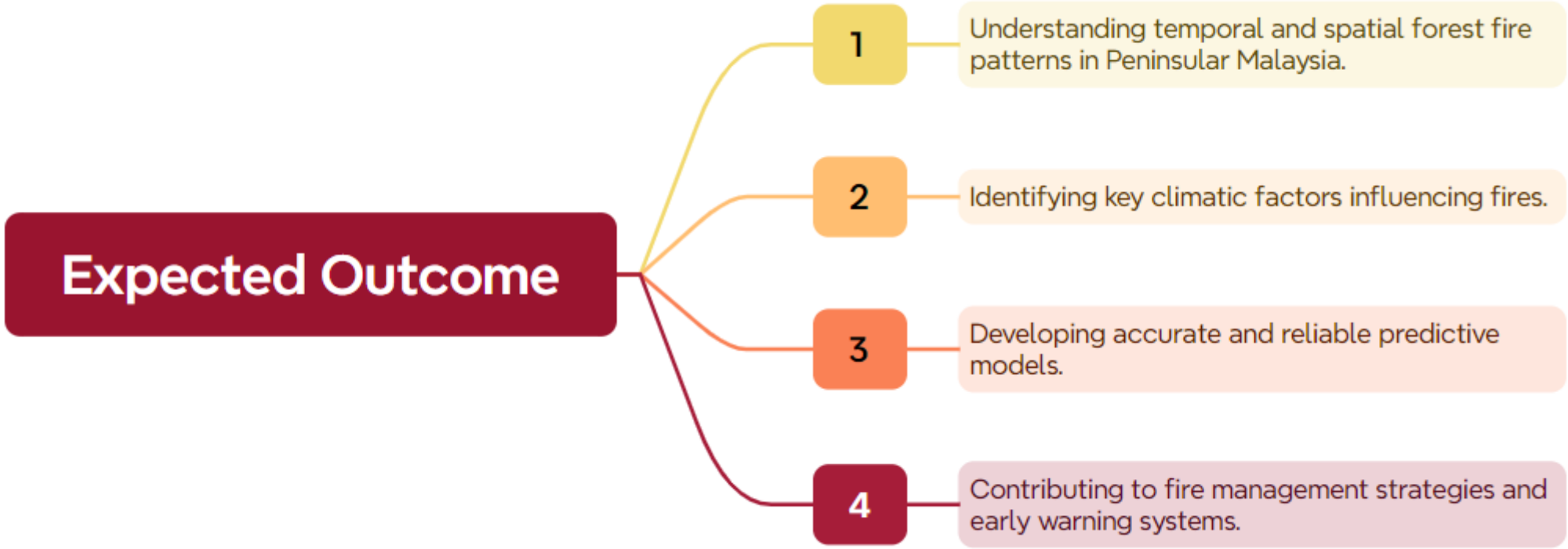
INITIAL FINDINGS

FEATURE ENGINEERING



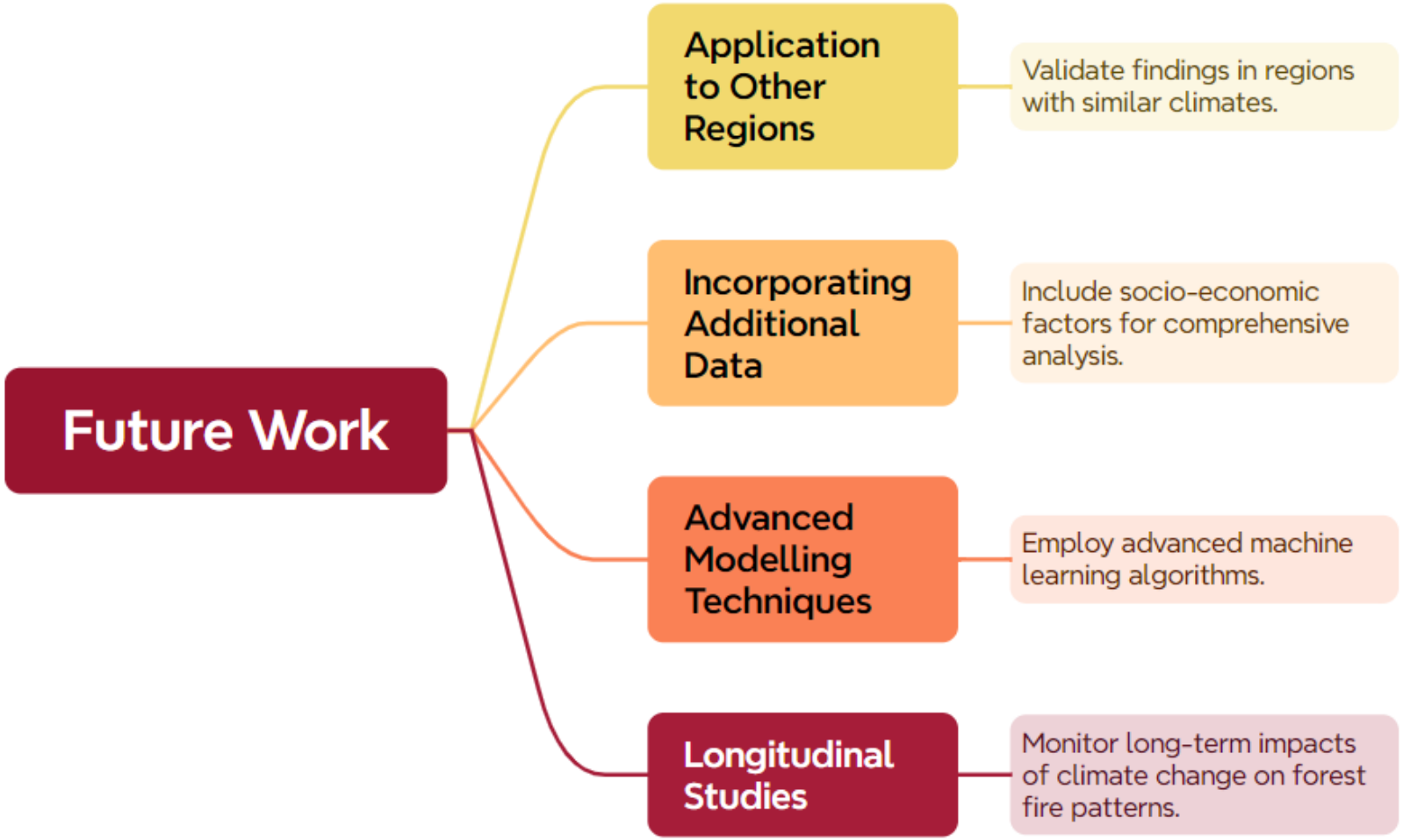
INITIAL FINDINGS

EXPECTED OUTCOME



INITIAL FINDINGS

FUTURE WORK



UTM
UNIVERSITI TEKNOLOGI MALAYSIA

univteknologimalaysia utm_my utmofficial

Thank You

www.utm.my

innovative • entrepreneurial • global