PREDICTION BASED WORKLOAD PREFORMACE EVALUATION OF LARGE SPATIAL DATABASE FOR DSS BASED DISASTER MANAGEMENT

1)Nanna Suryana Herman 2)Muhammad Syaifur Rohman,3)Fandy Setyo Utomo

1) Universiti Teknikal Malaysia Melaka
2) Universitas Dian Nuswantoro
3) STMIK Amikom Purwokerto

Volcano Eruption

“TERRIFYING NATURAL DISASTER FOR ALL MANDKIND IS VOLCANO ERUPTION”

Merapi volcano has given big many disadvantage and also can be categorize with biggest eruption on the world. Merapi volcano can be categorize very dangerous because according to modern research records had having eruption every two to five years and is surrounded by a very dense settlement. Since 1548, the volcano has erupted 68 times.

Background

- Issue come up after disaster occur there is need to make workload performance efficient and effective to support any response for Disaster Management
- There is need workload performance that can handle large spatial database

Problem Statement

- In order to achieve efficient and effective workload performance this research propose Workload Prediction for Large Spatial Databases.
- Careful selection of knowledge, algorithm and approach to support this study

Research Question

- How to give effectiveness and efficiently prediction of workload prediction to large spatial databases.
- How to give prediction when incoming workload are DSS, OLTP, or MIX of them.

Objective

- To proposed effectiveness and efficiently workload prediction technique to large spatial databases system.
- To automatically predicting a DBMS workload as either OLTP, DSS or MIX.

Research Contribution

- An approach of Pruning Decision Tree for large spatial database.
- An algorithm of Workload Prediction with Pruning Decision Tree method.

Literature Review

Refugee camp disaster management
GIS and TOPSIS model shows that suggested alternative camp locations

ÇETINKAYA, ÖZCEYLAN, ERBAŞ, & KABAK, 2016

Long, Liu, & Luo, 2008

Response and Evacuation Plan
Combination of B/S and C/S by VS2005 for reduction accidents management

Flood and landslide management
collaborative and decision making process for flood and landslide management.

Aye et al., 2016

Disaster Management

Hashemi et al., 2013

Seismic disaster prediction
Pre and post disaster response planning, land use and town planning, and risk management

Flood disaster prevention
Combinations of GIS application, flood hazard and risk scenarios for small scale

Lepuschitz, 2015

Uchida et al., 2013)

Earthquake disaster management
DTN with CWN (Cognitive Wireless Network) for Disaster Information Network System

LARGE SPATIAL DATABASE

Large Repositories of databases

Hariharan et al., 2007
SK queries with novel indexing data structure called KR* tree for processing SK queries

Large spatial and temporal datasets

L. Wang & Cheng, 2007
SDSS implemented for on-line flood forecasting and flood risk mapping disaster

Large scale crowd evacuation

Zhan, Lo, & Yuen, 2010
GIS implemented by urban emergency decision support model for large scale crowd evacuation

Large Urban Geological Information

Case & City, 2011
The research analyzed the requirements of geological safety supporting system for mega city regions

Searching in Large Scale disaster

Hayashi et al., 2015
Search method that searches a database storing many scenarios of disaster simulation

Workload Management

Decision Tree (Elnaffar et al., 2007)

Is it DSS or OLTP: automatically identifying DBMS workloads

Issue in Large Data

Discrete Fourier Transform (Holze et al., 2010)

Towards Workload Aware Self-Management: Predicting Significant Workload Shifts.

History does not match with response DM

Decision tree (Savage & Williams, 2012)

Performance Improvement in DBMS

Issue in Large Data

Fuzzy Logic (Abdul et al., 2014)

Database workload management through CBR and fuzzy based characterization

Membership function issues

RESEARCH DESIGN

DEFINE PROBLEM STATEMENT

Workload Prediction on Large Spatial Database

APROACH

Workload Data

ARCHITECTURE DEVELOPMENT

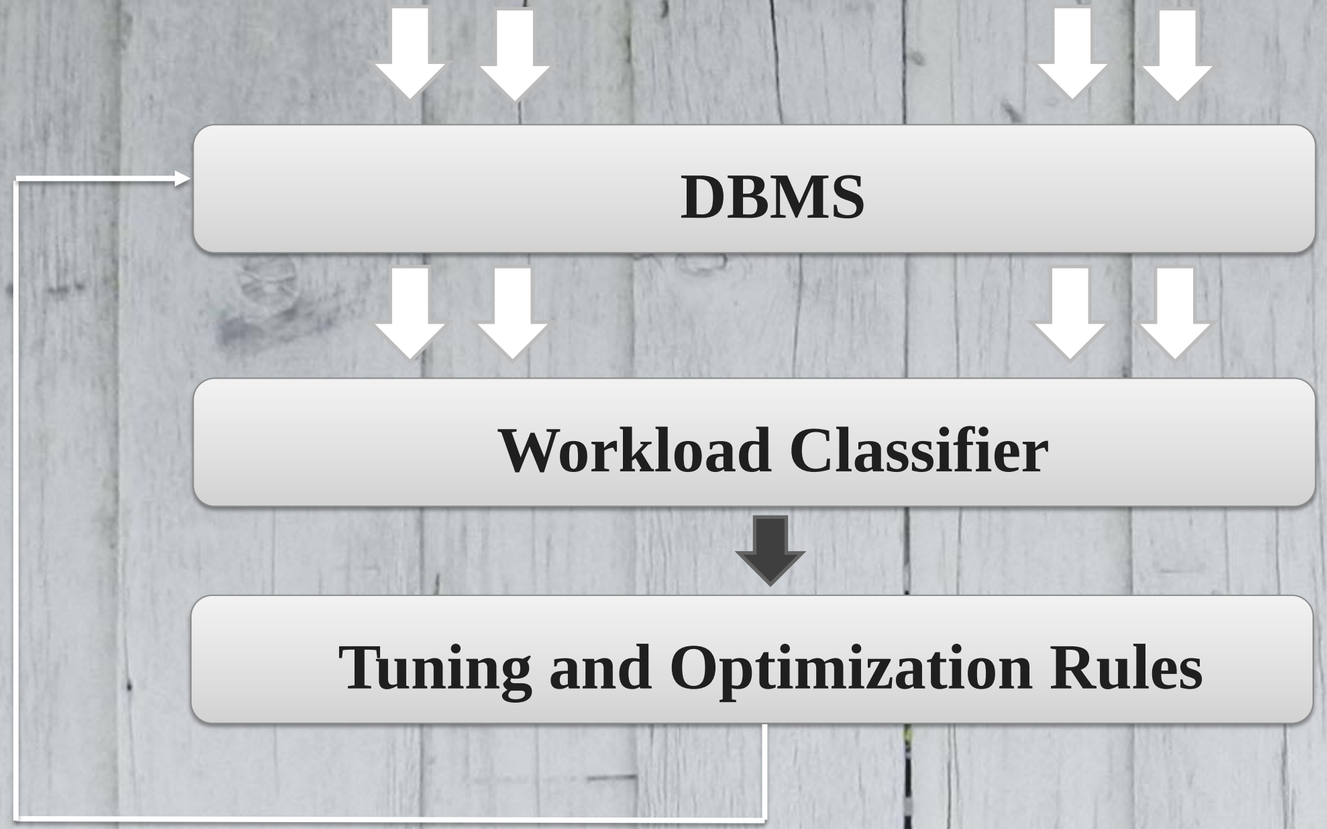
Workload Classifier | Tuning and Optimization Rules

IMPLEMENTATION

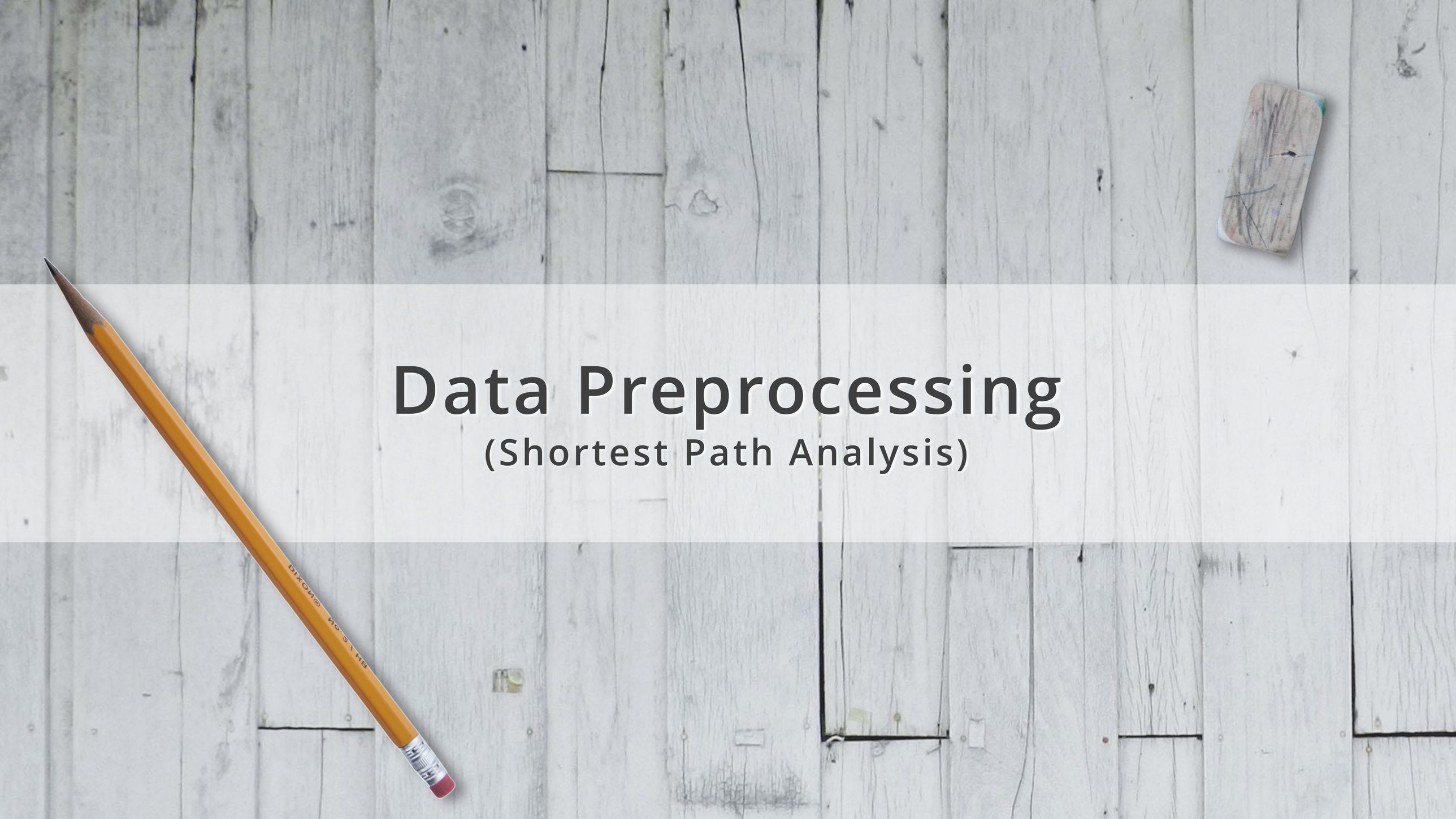
Decision Tree | Post Pruning

KNOWLEDGE

Identified Workload | Predict Incoming Workload



Architecture Development



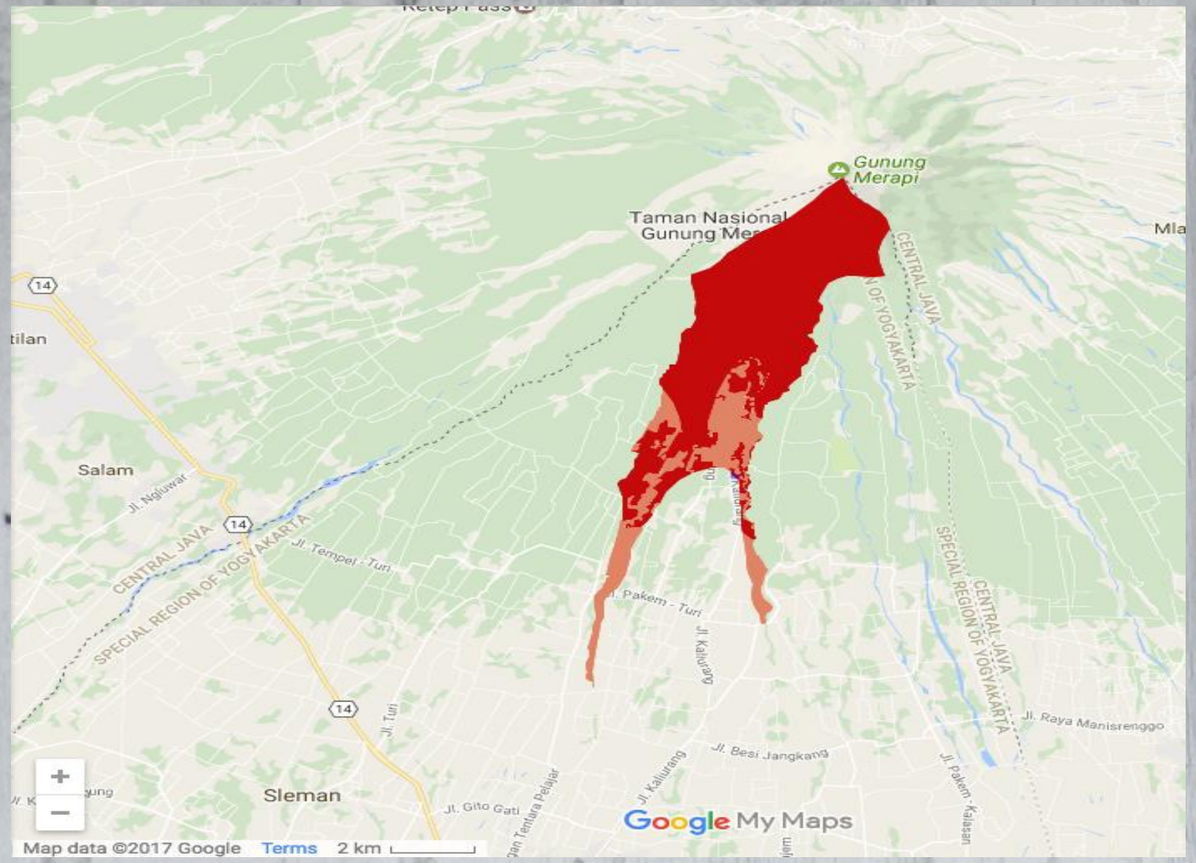
Data Preprocessing

(Shortest Path Analysis)

Configuration

Some configurations that have been involved to make sure all processes able to run precisely

Data Gathering



Disaster Area

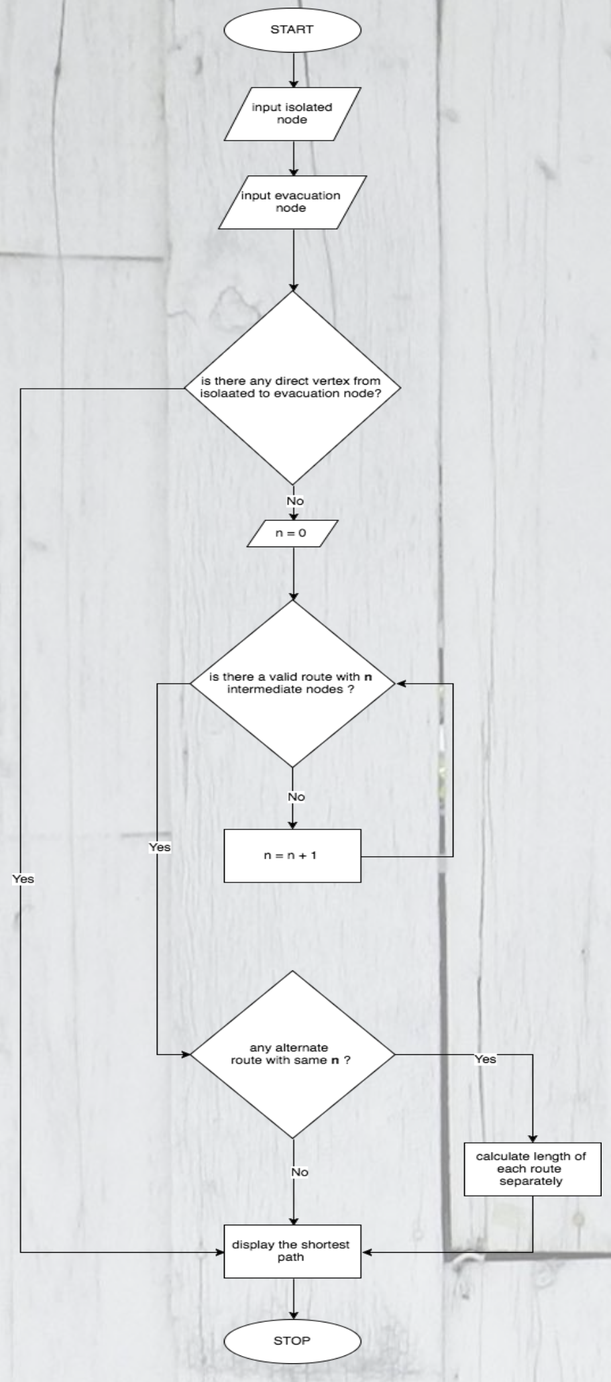
OpenStreetMap (OSM)

Node Disaster

Isolated Area

Evacuation Point

Dijkstra Flowchart



Dijkstra Algorithm Implementation

Route	Isolated Area Node	Nearest Source Node	Evacuation Node (Nearest 750)	Distance
1.	752	35	1	10.17 km
2.		47	1	10.47 km
3.	753	83	1	7.89 km
4.		620	1	9.07 km
5.	754	105	1	7.49 km
6.		118	1	7.13 km
7.	755	4	1	7.34 km
8.		417	1	7.48 km

Route Generated

Route 1

Distance 10.17 km

Route 2

Distance 10.47 km

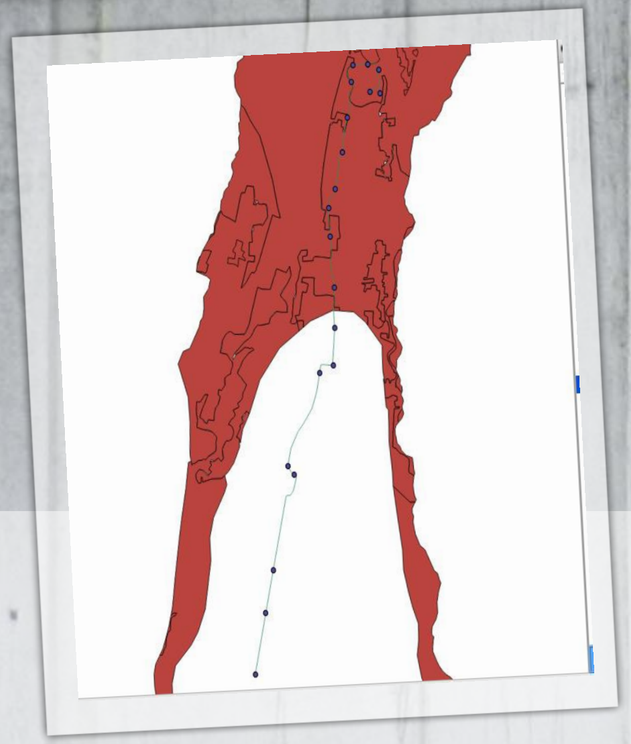
Route 3

Distance 7.89 km

Route 4

Distance 9.07 km

Route Generated



Route 5

Distance 7.49 km

Route 6

Distance 7.13 km

Route 7

Distance 7.34 km

Route 8

Distance 7.48 km

Workload Data Analysis

OLTP, DSS or MIX

Selection

2423		INSERT INTO result SELECT * FROM one619;
2424		SELECT * FROM result;
2425		SELECT * FROM result;
2426		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 552;
2427		SELECT * FROM result;
2428		SELECT * FROM result;
2429		DELETE FROM result;
2430		SELECT * FROM result;
2431		SELECT * FROM result;
2432		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 68;
2433		DELETE FROM result;
2434		DELETE FROM result;
2435		SELECT * FROM result;
2436		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 215;
2437		DELETE FROM result;
2438		SELECT * FROM result;
2439		SELECT a.*, ST_AsText(b.geom) FROM pgr_dijkstra('SELECT gid AS id, source, target, cost FROM spatial_merapi', 93, ARRAY[401,402,403,404,405,1,406,407], FALSE) AS a LEFT JOIN spatial_merapi as b ON (a.edge = b.gid) ORDER BY seq;
2440		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 347;
2441		SELECT * FROM result;
2442		SELECT * FROM result;
2443		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 80;
2444		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 569;
2445		DELETE FROM result;
2446		DELETE FROM result;
2447		UPDATE spatial_merapi SET cost = 10000 WHERE gid = 134;
2448		DELETE FROM result;

workload data exist only in this study was workload that involve in the process of shortest path analysis using Dijkstra algorithm for evacuation from isolated area to evacuation point.

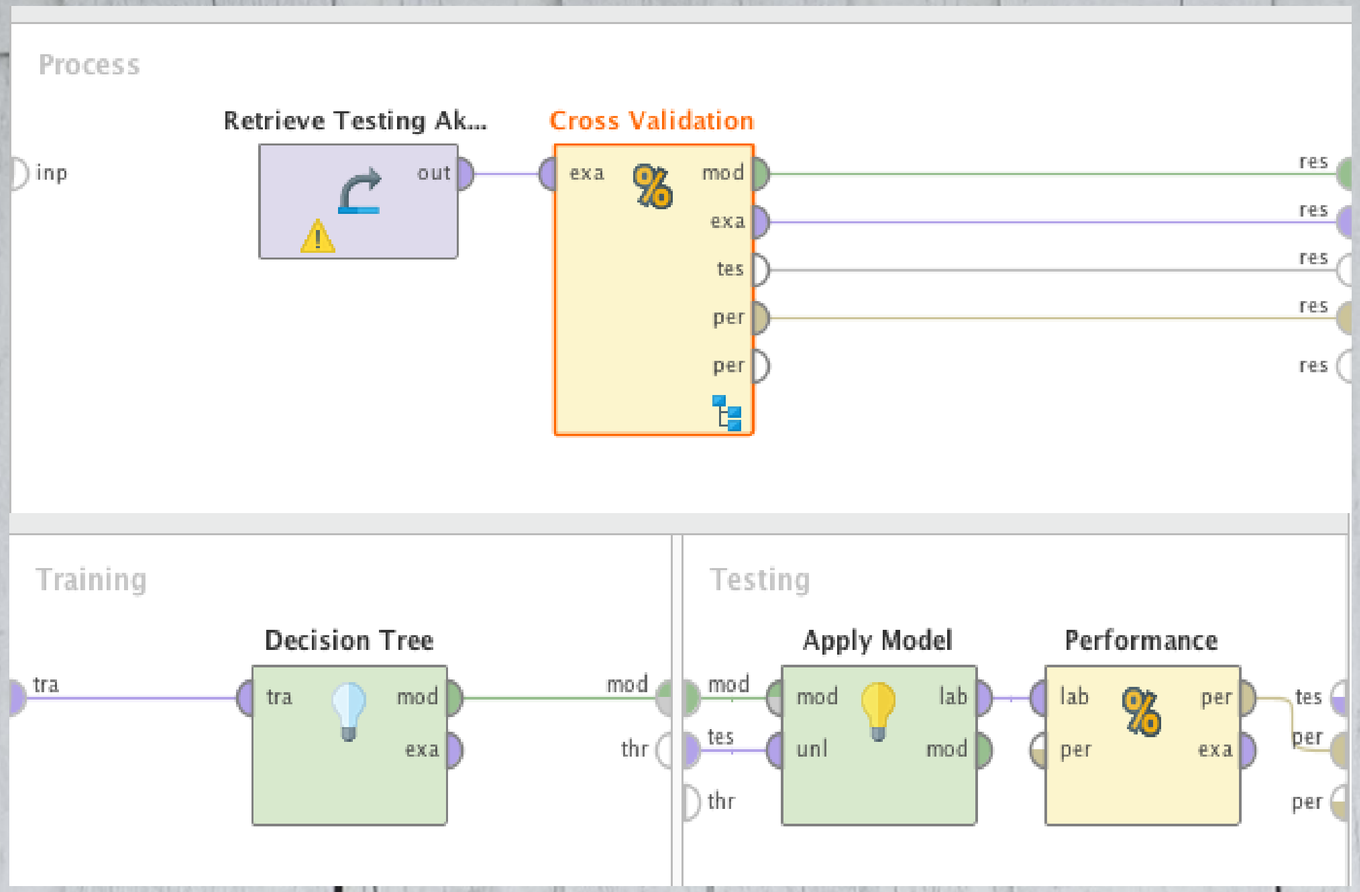
Data Transformation

Data transformation is performed by changing some of data to make data mining process become easier after the Example set are selected. Besides that, transformation process also improves the accuracy and algorithm efficiency.

	A	B	C	D	E
1	select	update	delete	insert	label
2	1	0	0	0	DSS
3	1	0	0	1	MIX
4	0	1	0	0	OLTP
5	0	0	1	0	OLTP
6	1	0	0	0	DSS
7	0	0	1	0	OLTP
8	1	0	0	0	DSS
9	0	0	1	0	OLTP
10	0	0	1	0	OLTP
11	1	0	0	0	DSS
12	1	0	0	0	DSS
13	0	1	0	0	OLTP
14	1	0	0	0	DSS
15	0	0	1	0	OLTP
16	0	1	0	0	OLTP
17	0	0	1	0	OLTP
18	0	0	1	0	OLTP
19	1	0	0	0	DSS
20	0	0	1	0	OLTP
21	0	0	1	0	OLTP
22	0	1	0	0	OLTP
23	1	0	0	1	MIX
24	2	0	0	0	MIX
25	0	0	1	0	OLTP
26	0	1	0	0	OLTP
27	0	1	0	0	OLTP
28	0	0	1	0	OLTP
29	0	0	1	0	OLTP
30	1	0	0	0	DSS
31	0	1	0	0	OLTP
32	1	0	0	0	DSS
33	0	0	1	0	OLTP
34	1	0	0	0	DSS
35	2	0	0	0	MIX
36	0	1	0	0	OLTP
37	0	1	0	0	OLTP
38	1	0	0	0	DSS
39	0	0	1	0	OLTP
40	0	1	0	0	OLTP
41	0	1	0	0	OLTP

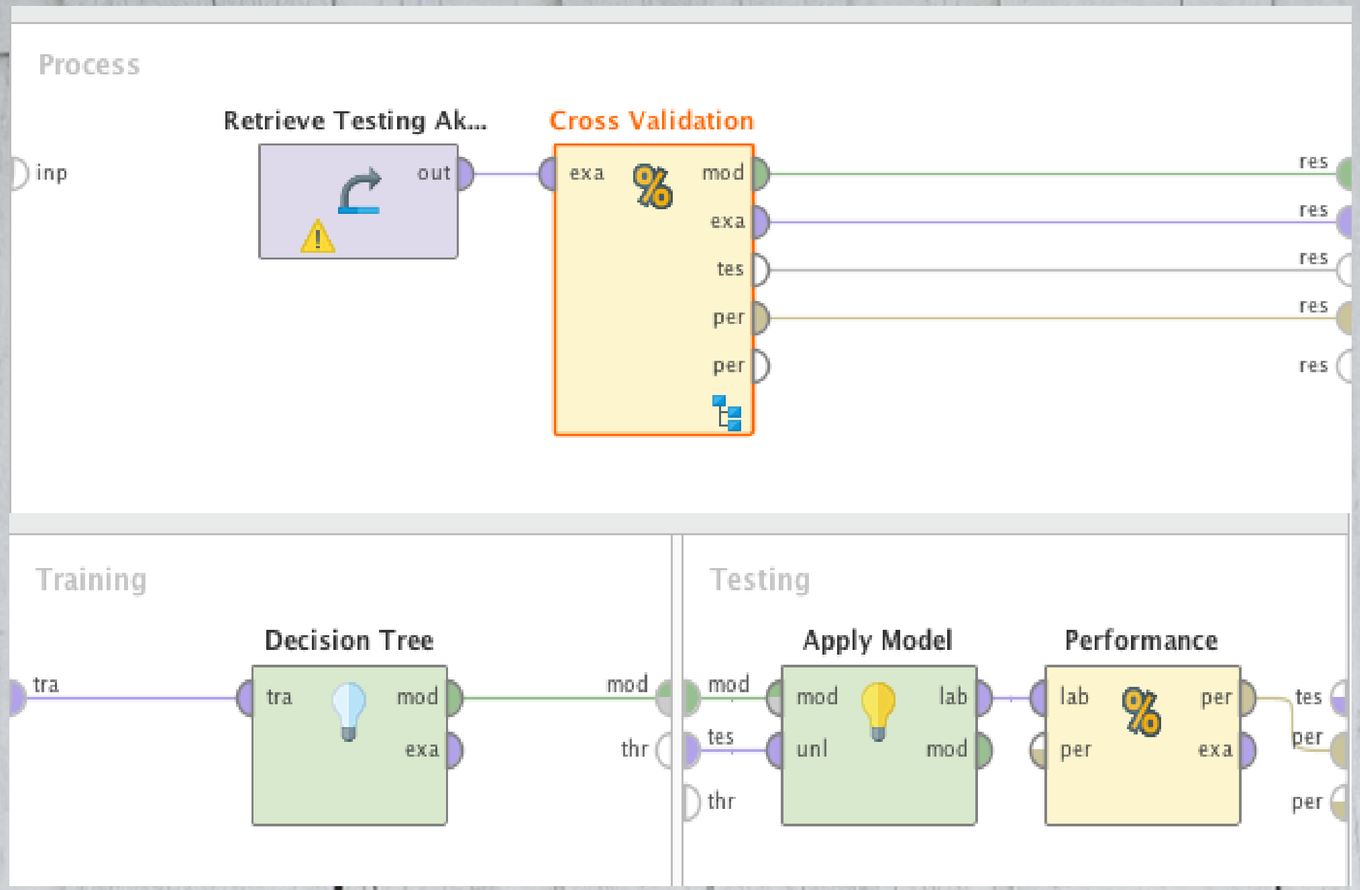
Validation and evaluation

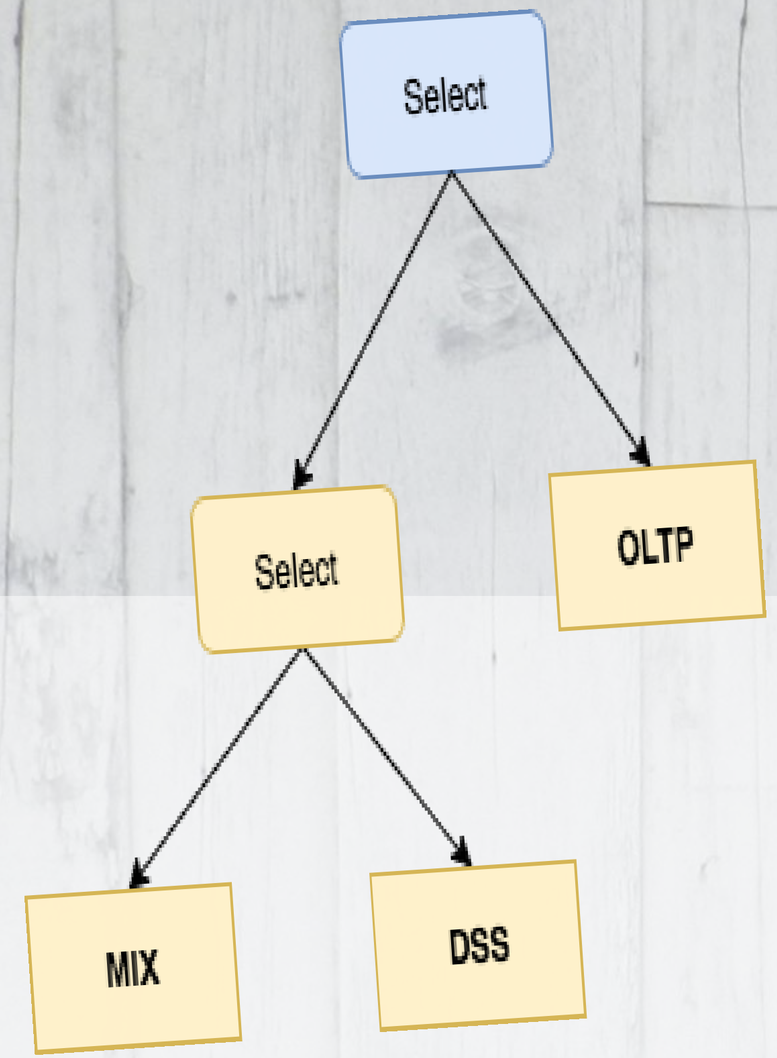
In this phase, accuracy measuring from Decision Tree Algorithm using Cross Validation method was implemented.



Testing and result

The best result of testing for accuracy measuring of Decision Tree algorithm using Workload data and cross validation with 10 folds provided in RapidMiner framework shown in confusion matrix image beside.





RESULT

After the example set ready to be proceed, the testing phase was started. The best testing result shown the result of accuracy that could categorize as excellent classification.

Conclusion

- The analytical of approach of Pruning Decision Tree for large spatial database with flexible capability improvement especially for evacuation phase of DM.
- The enhancement and development of Algorithm of Workload Prediction with Pruning Decision Tree method presented to improve the capability.

THANK YOU!

Do You Have Any Questions?