

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines



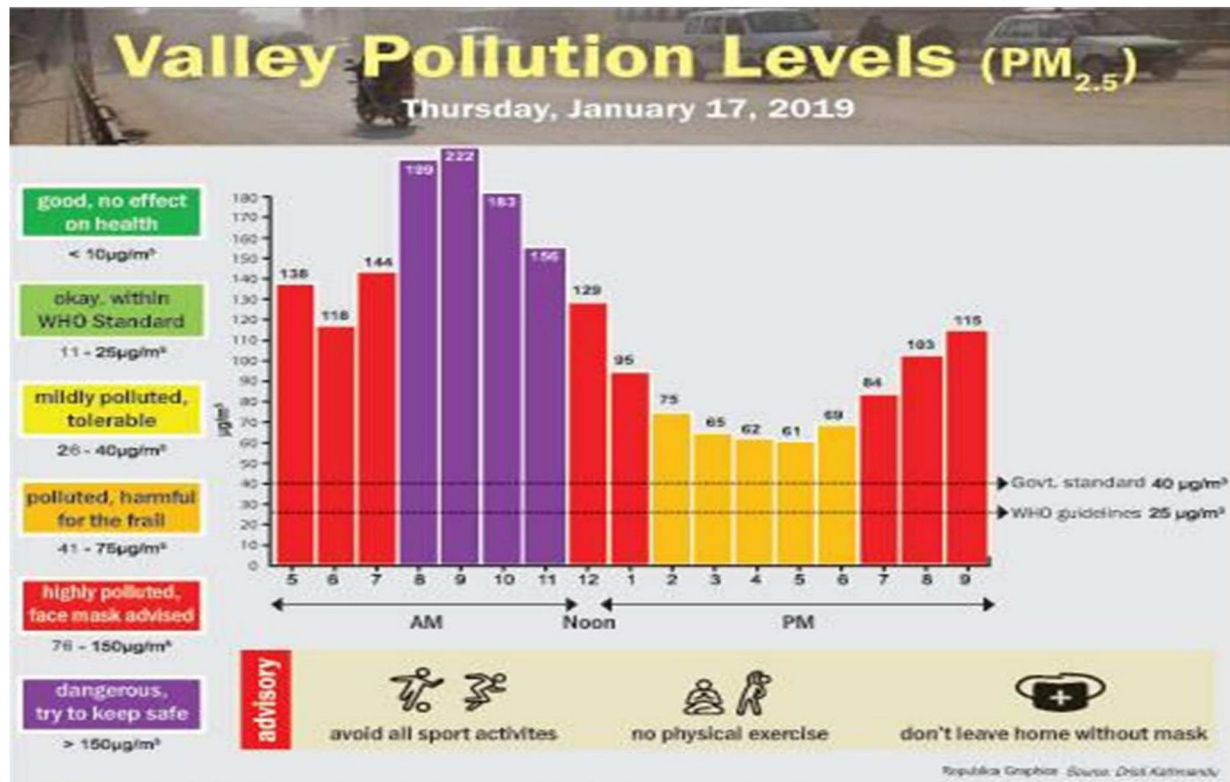
Supervised by :
Dr Nanda Bikram Adhikari

UGC Collaborative Research Grant
(Award, CRG-73/74-01Egg)

Alina Devkota 
Saloni Shikha 
Spandan Pyakurel 
Sushant Gautam 



Motivation



Republica

Page Liked · 13 hrs ·

Kathmandu's Pollution Level at alarming levels.

Advisory: Avoid All Sports Activities, No Physical Exercise (especially during morning), and Don't leave home without a mask.

#nepal #pollution #kathmandu #health #AirPollution #children #women #elderly

👍 15

6 Comments 19 Shares



Like



Comment



Share

Most Relevant ▾



Saugat Rai kati hernu yo report ?? jaile dekhakai ta ho, bhogekai ta ho.



Write a comment...



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Kathmandu



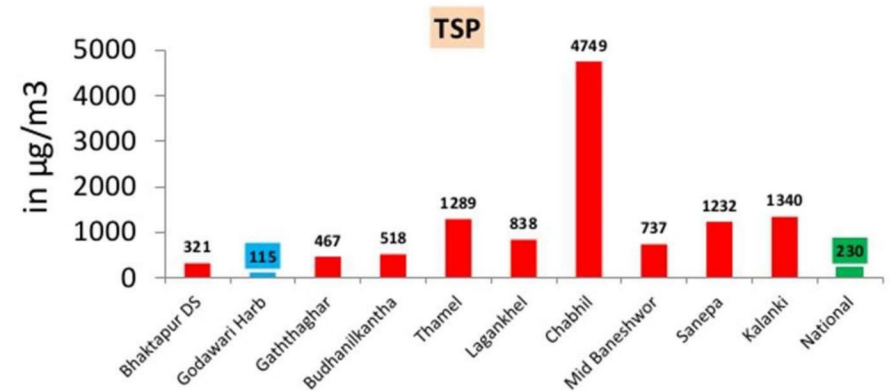
Second most polluted city
in the world

Most polluted city in Asia



Source: 2018 Mid-Year Pollution Index

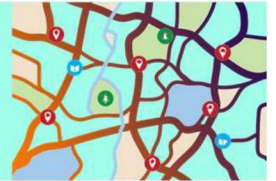
Air Quality of Nepal/Kathmandu Valley



Air Quality Trend in Kathmandu valley (24-hr average): May, 2017 (Source : Quest Forum/MoPE Report, 2017)

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Building Pollution Profile of the City



DOECE Server



Access
data 
and



powerful analytics tools

from

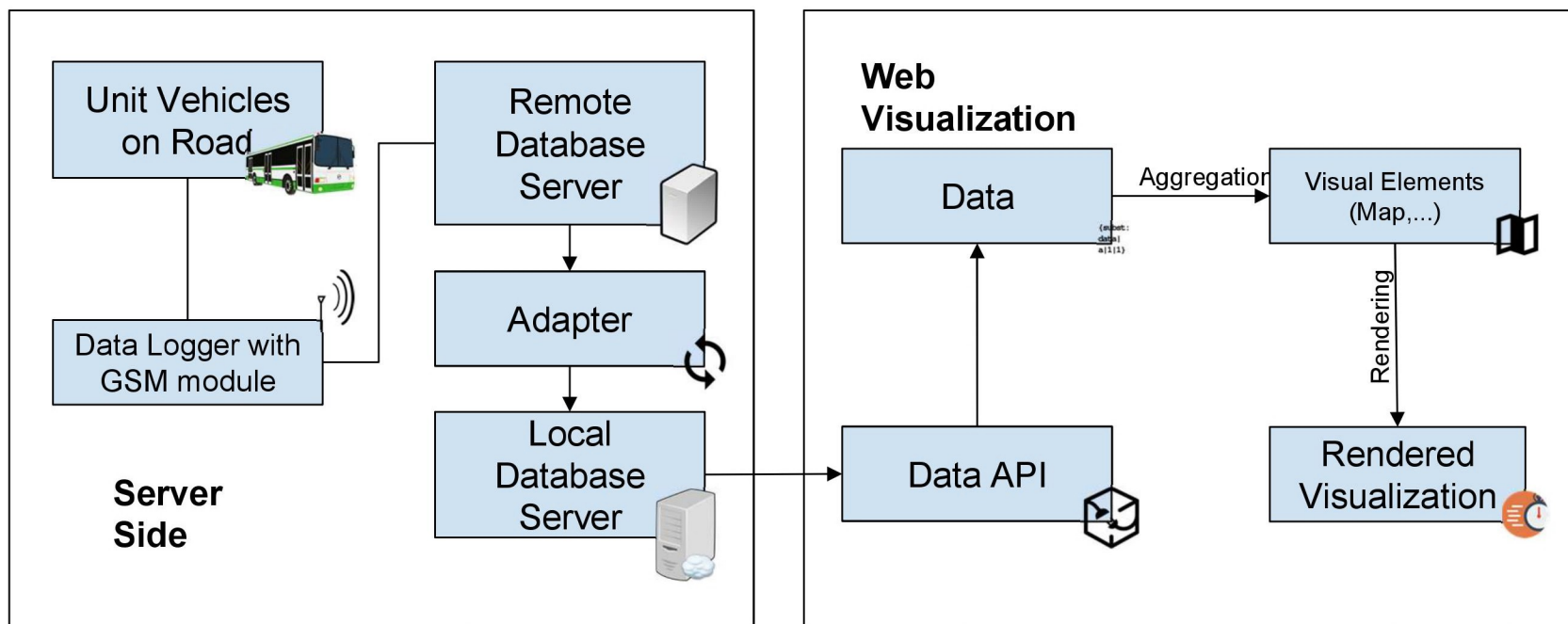
client terminal 

through

 **Internet.**



Working Principle



Bus Mount Sensor



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Off-Road Data Collection



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Pollution Sensor



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Measures



Weather

Temperature(C)	Rain Rate(mm/h)
Humidity(%)	Daily Rain(mm)
ABS Pressure(mmHg)	Weekly Rain(mm)
REL Pressure(mmHg)	Monthly Rain(mm)
Wind(km/h)	Yearly Rain(mm)
Gust(km/h)	Solar Rad(lux)
Dew Point(C)	UV(uW/cm2)
Wind Direction(deg)	UVI()



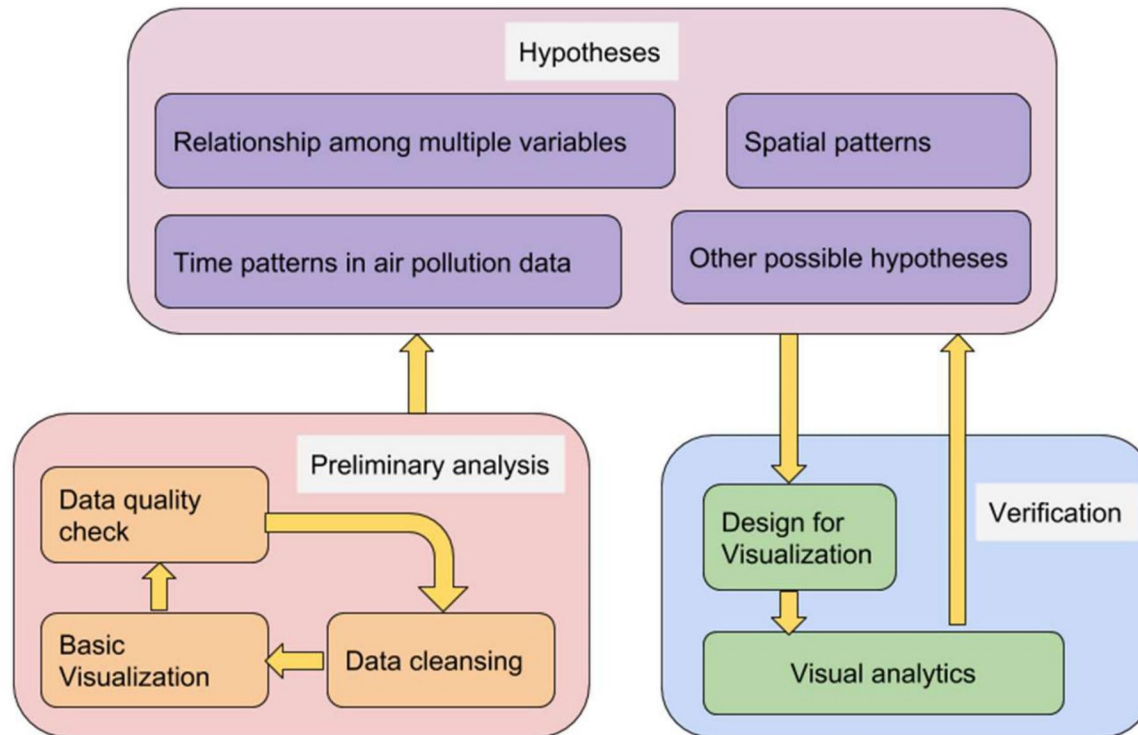
Pollution

Date Time
PM2.5
CO2
Formaldehyde
VOCs

Along with GPS coordinates and Timestamp




Visualization Workflow for Air Pollution Data



Preprocessing:

1. Remove low-quality columns
2. Replace missing values
3. Removing >0.8 Correlated columns: Outdoor Temp, Outdoor Humidity and Daily Rain
4. Changing: Date to Hours of Day

Output: 25 down to 9 attributes



Time Number	Indoor Temperat... Number	Outdoor Temper... Number	Dew Point(C) Number	ABS Pressure(hpa) Number	Wind Direction(d... Number	Rain Event(mm) Number	Weekly Rain(mm) Number	Solar Rad(w/m2) Number
19	20.300	17	14.100	870.200	107	0	0	0
19	20.300	17	14.100	870.200	107	0	0	0
19	20.300	17	14.100	870.400	107	0	0	0
19	20.300	17	14.100	870.600	107	0	0	0
19	20.300	17	14.100	870.400	107	0	0	0
19	20.300	17	14.100	870.500	107	0	0	0
19	20.300	17	14.100	870.700	107	0	0	0
19	20.300	17	14.100	870.600	107	0	0	0
19	20.300	17	14.100	870.100	107	0	0	0
19	20.300	17	14.100	870.200	107	0	0	0
19	20.200	17	14.100	870.200	107	0	0	0
19	20.300	17	14.100	870.100	107	0	0	0
19	20.300	17	14.100	870	107	0	0	0

5,385 rows - 9 columns (9 numerical)

Data Sample Analysis



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

1. Clustering

k-Means - Summary

Number of Clusters: 2
Distance Measure: Squared Euclidean Distance
Average Cluster Distance: 4.315
Davies-Bouldin Index: 1.018

Cluster 0

242

Average Distance: 4.821

Indoor Temperature(C) is on average **128.54%** larger, **Weekly Rain(mm)** is on average **54.67%** larger, **Outdoor Temperature(C)** is on average **46.44%** larger

Cluster 1

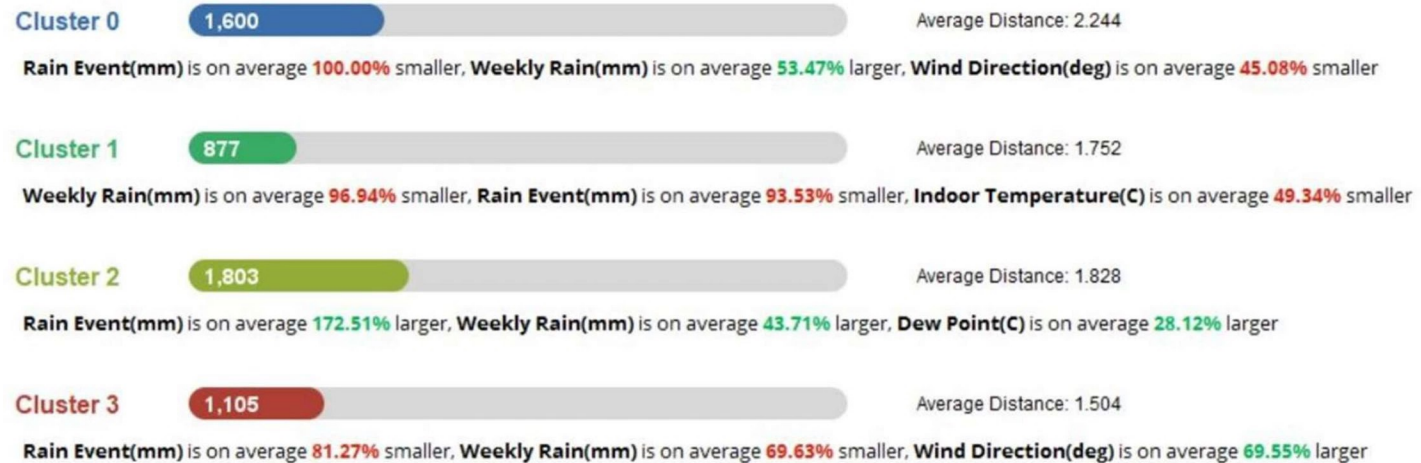
5,143

Average Distance: 4.291

Indoor Temperature(C) is on average **6.05%** smaller, **Weekly Rain(mm)** is on average **2.57%** smaller, **Outdoor Temperature(C)** is on average **2.19%** smaller

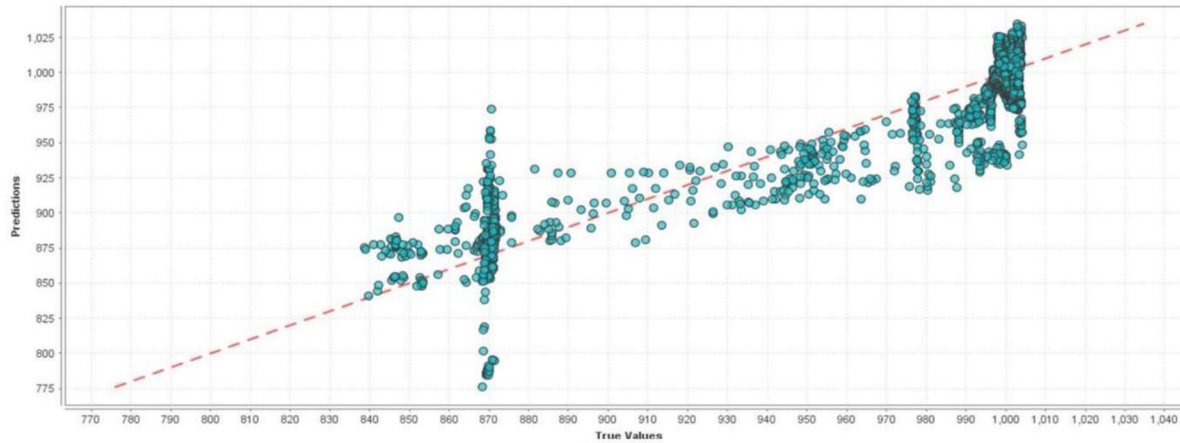
x-Means - Summary

Number of Clusters: 4
Distance Measure: Euclidean Distance
Average Cluster Distance: 1.873
Davies-Bouldin Index: 0.743

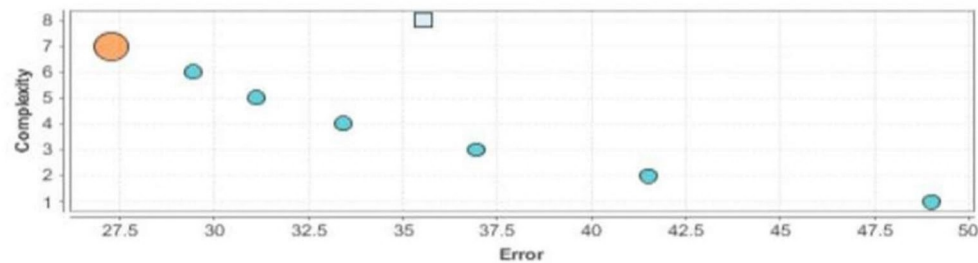


Regression Modelling for ABS Pressure (hPa) using other 5 columns as input

Generalized Linear Model - Predictions Chart



Optimal Trade-offs between Complexity and Error



● Used feature set ● Optimal trade-offs □ Original feature set ● Shown below

DeepLearning

MSE: 180.24547

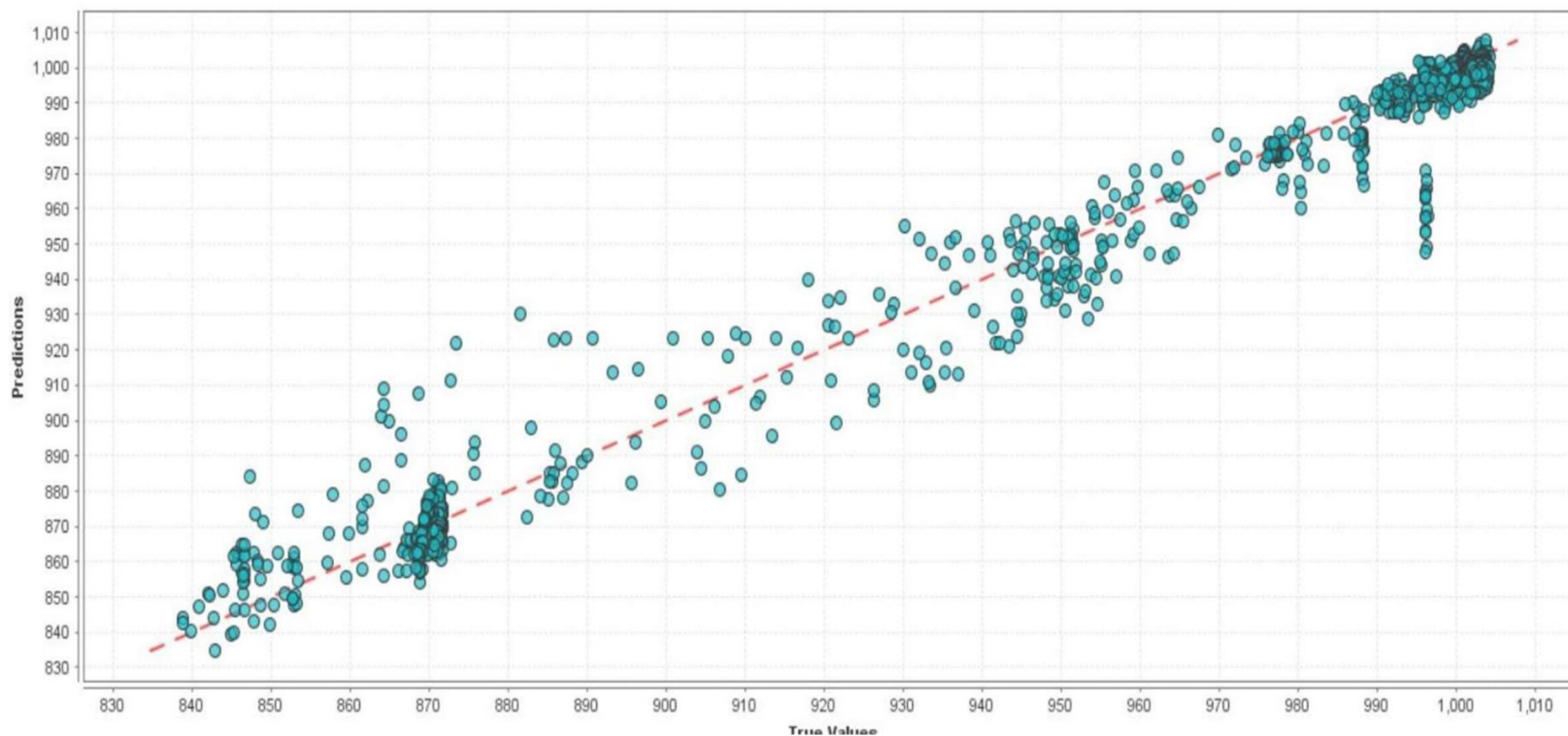
R^2: 0.9558107

Mean residual deviance: 180.24547

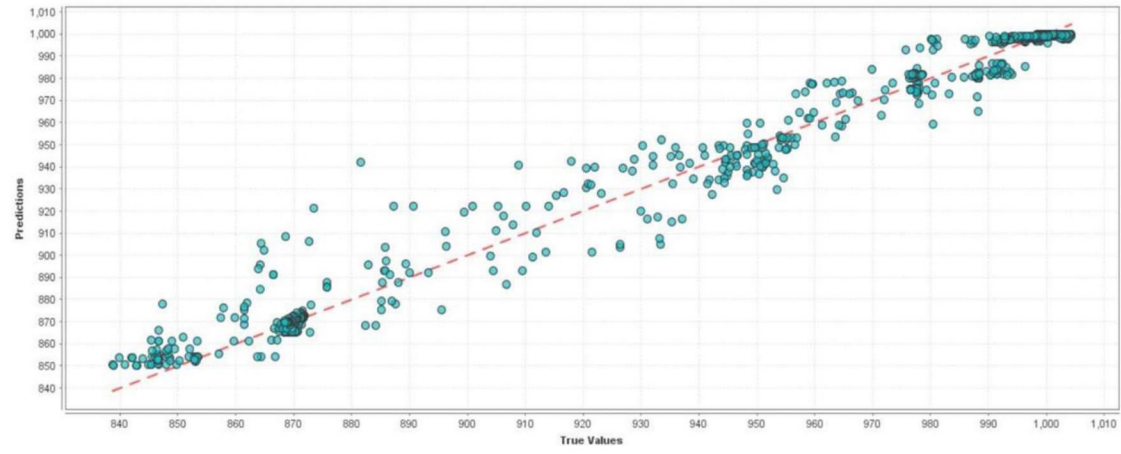
Scoring History:

Timestamp	Duration	Training Speed	Epochs	Iterations	Samples	Training MSE	Training Deviance	Training R^2
2019-04-16 09:54:27	0.000 sec		0.00000	0	0.000000		NaN	NaN
2019-04-16 09:54:27	0.170 sec	22594 rows/sec	1.00000		1 3231.000000	315.05813	315.05813	0.92276
2019-04-16 09:54:27	0.332 sec	22833 rows/sec	2.00000		2 6462.000000	231.75128	231.75128	0.94318
2019-04-16 09:54:27	0.494 sec	22969 rows/sec	3.00000		3 9693.000000	210.15572	210.15572	0.94848
2019-04-16 09:54:27	0.654 sec	23161 rows/sec	4.00000		4 12924.000000	229.37043	229.37043	0.94377
2019-04-16 09:54:27	0.813 sec	23244 rows/sec	5.00000		5 16155.000000	198.20058	198.20058	0.95141
2019-04-16 09:54:28	0.971 sec	23328 rows/sec	6.00000		6 19386.000000	210.21262	210.21262	0.94846
2019-04-16 09:54:28	1.128 sec	23413 rows/sec	7.00000		7 22617.000000	182.34357	182.34357	0.95530
2019-04-16 09:54:28	1.286 sec	23476 rows/sec	8.00000		8 25848.000000	180.24547	180.24547	0.95581
2019-04-16 09:54:28	1.444 sec	23526 rows/sec	9.00000		9 29079.000000	214.38377	214.38377	0.94744
2019-04-16 09:54:28	1.599 sec	23618 rows/sec	10.00000		10 32310.000000	181.88087	181.88087	0.95541
2019-04-16 09:54:28	1.622 sec	23601 rows/sec	10.00000		10 32310.000000	180.24547	180.24547	0.95581

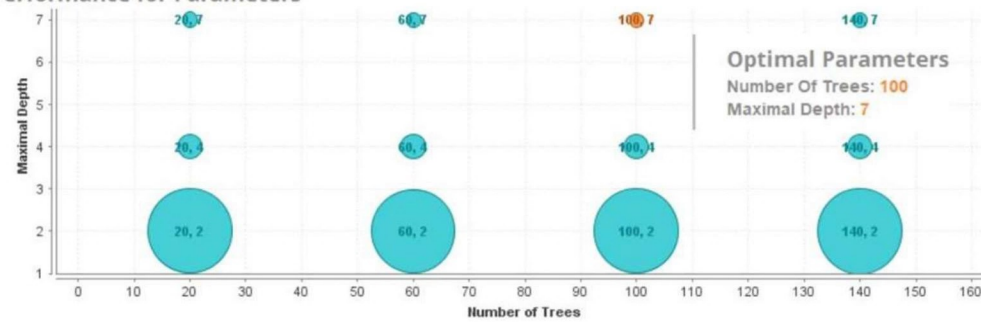
Deep Learning - Predictions Chart



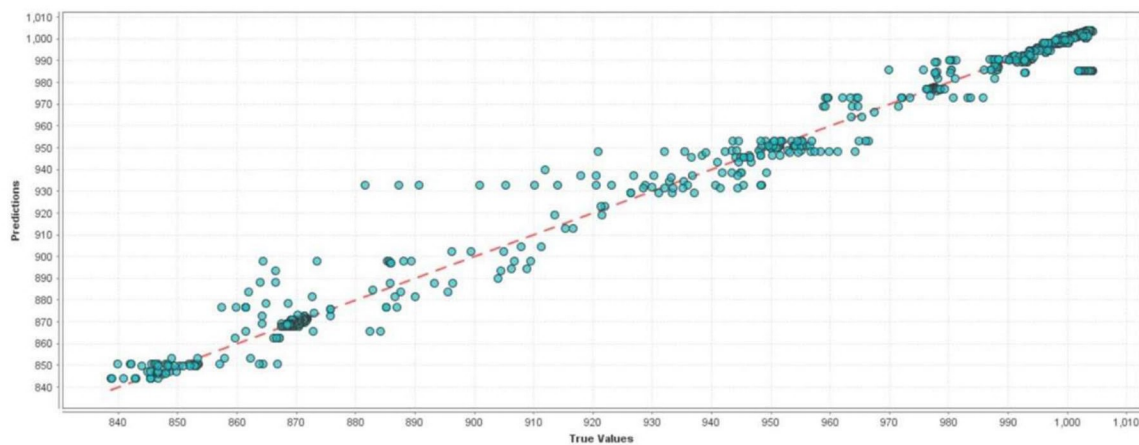
Random Forest - Predictions Chart



Performance for Parameters



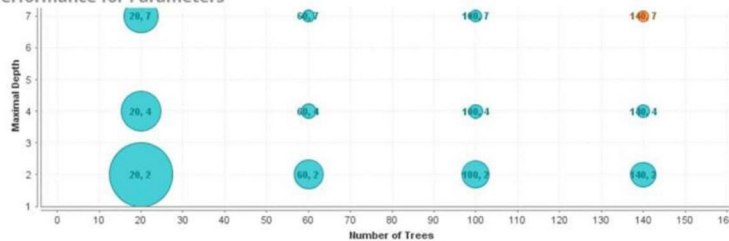
Gradient Boosted Trees - Predictions Chart



Optimal Parameters

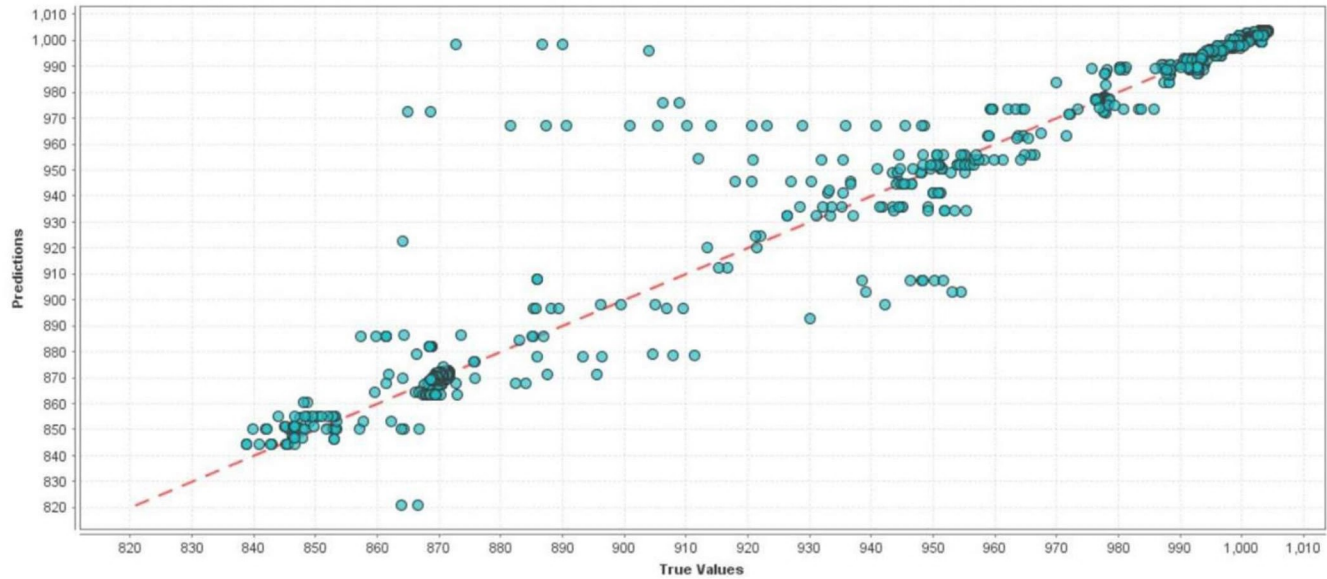
Number Of Trees: 140
Maximal Depth: 7

Performance for Parameters



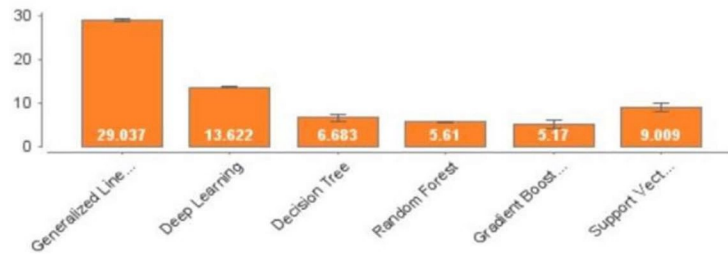
Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Support Vector Machine - Predictions Chart

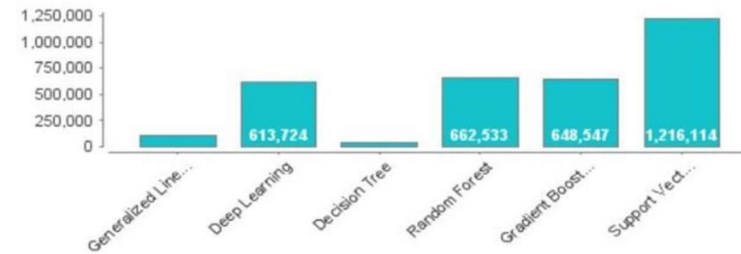


Although Low Data Range (Date)
 Tried **Regression** Modelling for **ABS** Pressure (hPa) using other 5 columns as input
 Indoor temp and Dew Point were highly correlated to **ABS** Pressure so not used.
 Automatic feature selection/generation technique used.

Root Mean Squared Error



Runtime (ms)

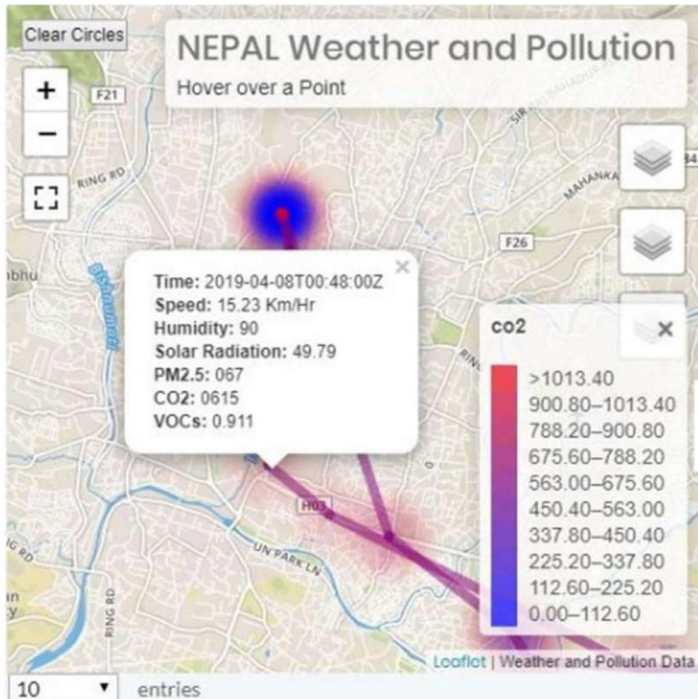


Model	Root Mean Squared Error	Standard Deviation	Runtime
Generalized Linear Model	29.0	0.2	109782.0
Deep Learning	13.6	0.2	613724.0
Decision Tree	6.7	0.8	35899.0
Random Forest	5.6	0.1	662533.0
Gradient Boosted Trees	5.2	1.0	648547.0
Support Vector Machine	9.0	0.9	1216114.0

Integrated Data Visualization Tool



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines



Weather and Pollution Data

Data from Mobile Vehicles



- id
- time
- latitude
- longitude
- speed
- distance
- humidity
- dew_point
- wind_direction
- wind_speed_mph_field
- wind_gust_mph_field
- rain
- daily_rain
- weekly_rain
- monthly_rain
- yearly_rain
- solar_radiation
- uv_index
- pressure
- pm2_5
- co2
- formaldehyde
- vocs

10 entries

Search:

time	latitude	longitude	speed	distance	humidity	dew_point	wind_direction	wind_speed_mph_field	wind_gust_mph_field	rain	daily_rain
2019-04-07T18:48:00Z	27.72181167	85.322665	0.47	0.000983	89	54	134	1.12	1.12	0	0
2019-04-07T20:23:00Z	27.72189	85.32245333	1.43	0.020182694	92	54	240	0.89	1.12	0	0

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Time Series Modelling using ARIMA

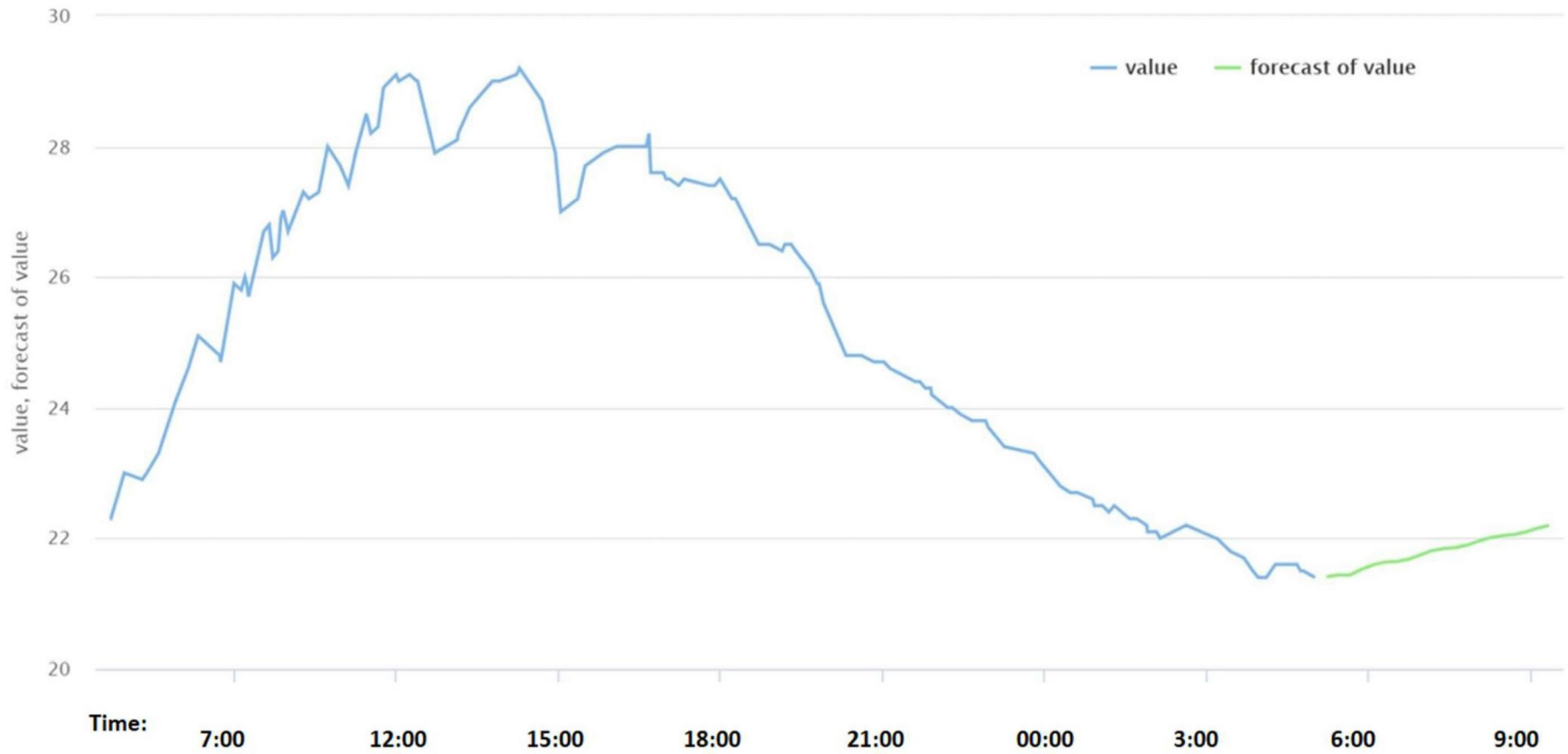


Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

General Overview

- An ARIMA model is a mathematical model for time series data.
- George Box and Gwilym Jenkins developed a systematic approach for fitting these models to data so these models are often called Box-Jenkins models.
- We always use statistical or forecasting programs to fit these models
 - The programs fit models and produce forecasts for us.

Pulchowk June 13, 2019 Data Forecast - Temperature



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

ARIMA MODEL

Arima Model trained on the following Time Series:
Number of values: 7605

Resulting Arima Model:

Arima Model (p: 4, d: 0, q: 5)

AR Coefficients: [1.4547171020898764,-0.16147656599579205,-0.1422626859811189,-
0.1547870960740363,]

MA Coefficients: [-0.4118651653385933,-0.23716908131826445,-
0.17587892444060524,0.04759654024745145,0.010011132295539817,]

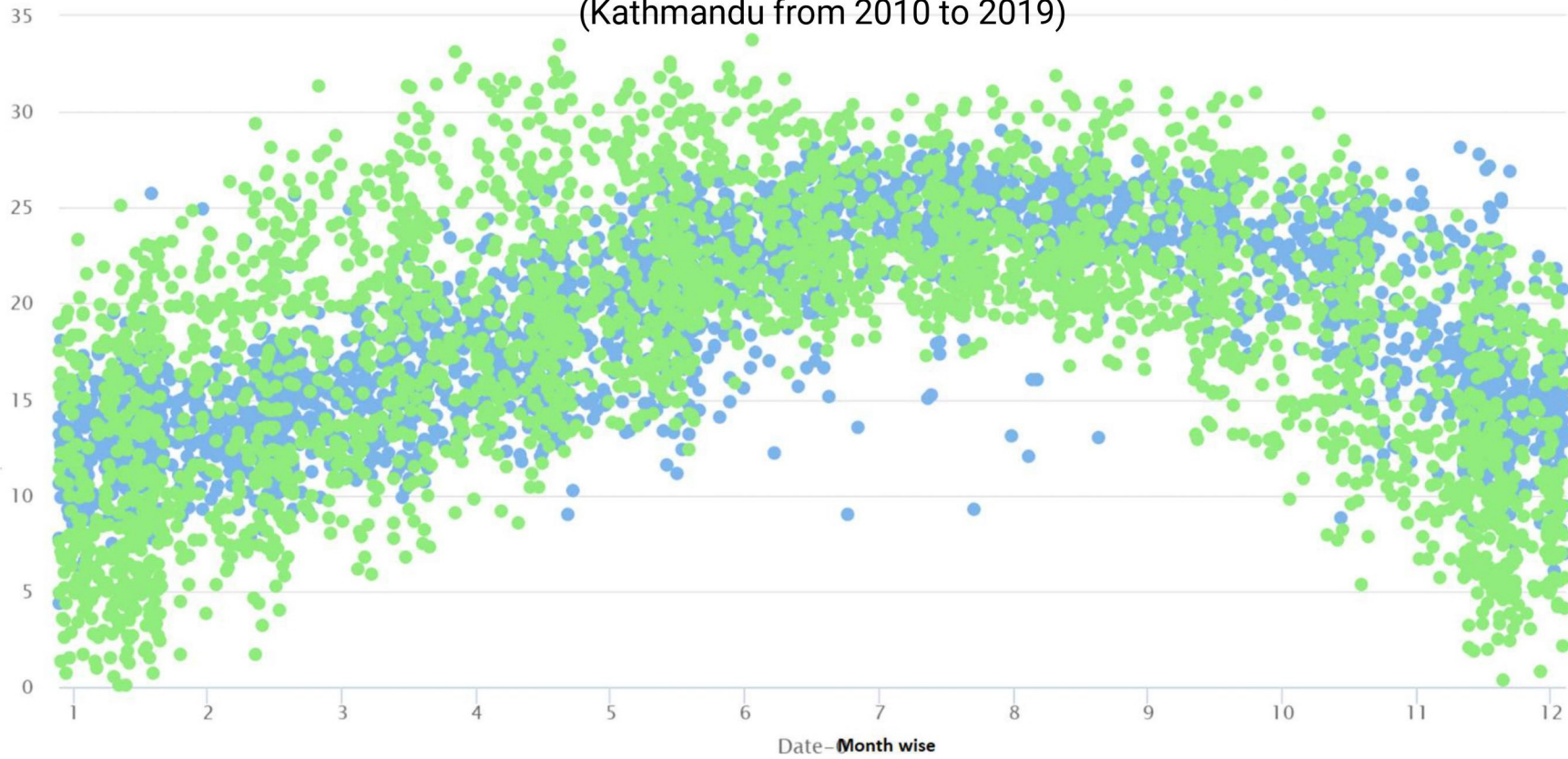
constant: 22.64487141982825

Time Series Modelling using RNN



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Prediction of Temperature variation in months relative to year using RNN (Kathmandu from 2010 to 2019)



● prediction ● label

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

RNN Model Parameters



Optimizer: adam
Loss Function: mean_squared_error
Activation: softplus

```
root_mean_squared_error: 5.760 +/- 0.000
```

Challenges



Massive Raw Data from Sensors

Unstable and Inconsistent Data

Unprocessed Spatial Data

Data Security

User Friendliness with AI Power

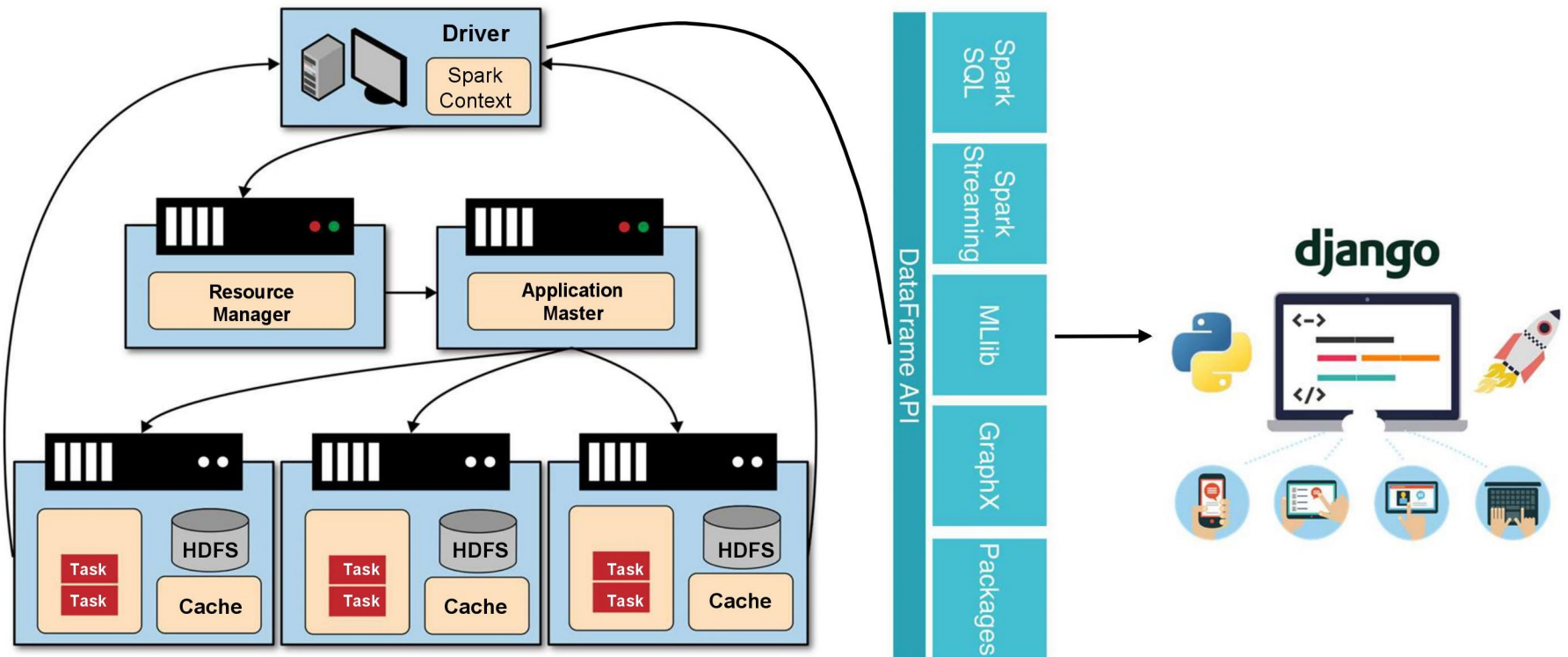
**Processing Power for
on-the-go model Training**

**Unstable and Inconsistent Data
Unprocessed Spatial Data**

**Availability of Responsive Server to
Handle Web Requests**

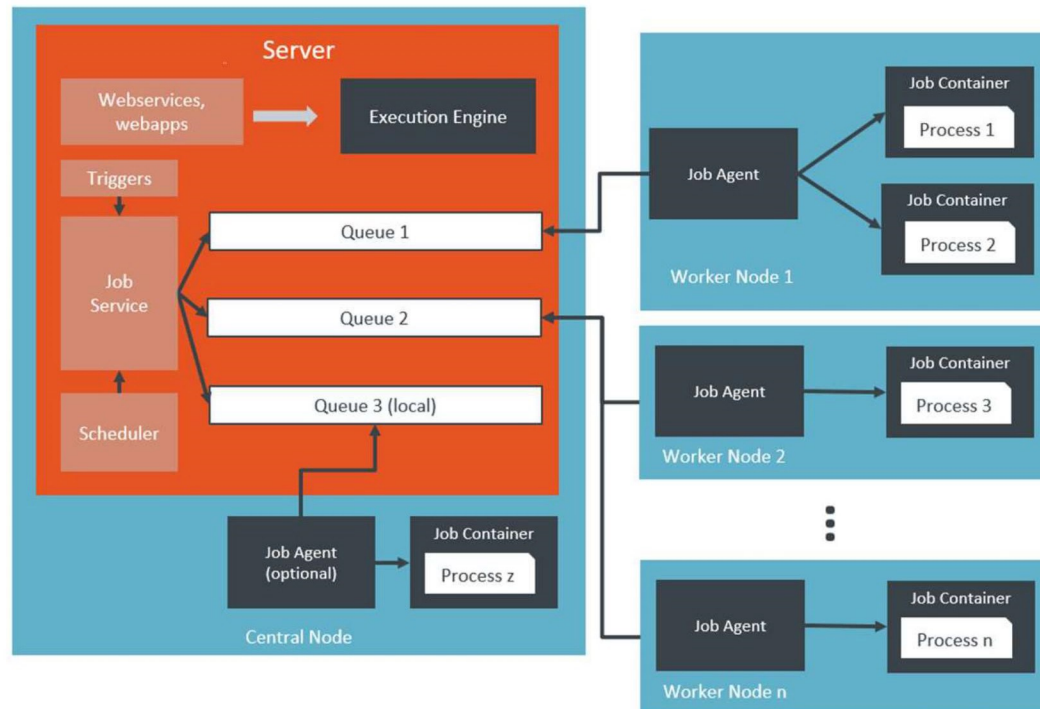
Integrity of Collected Data

Proposed Server Implementation at DOECE



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Server Architecture



Application



City Weather Map

City Pollution Map

Air Quality Modelling

Web based Visualization of Pollution
and Weather Proxies

Web based Visualization of
Pollution and Weather Proxies

Traffic Congestion Analysis

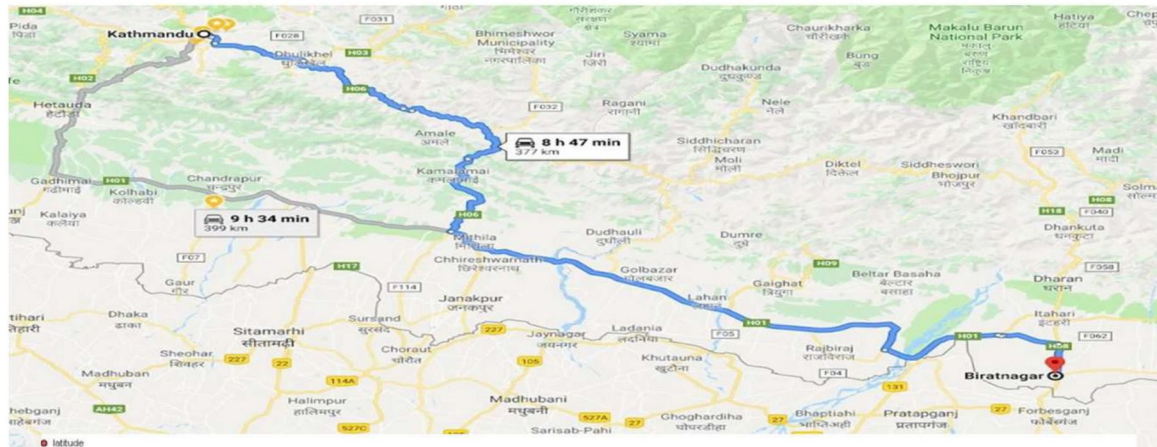
Bigdata Warehousing of Public
Remote Sensor Data

Routes from GPS Traces

Data Source for Further Researches

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

GPS 2019-04-08 to 11 Kathmandu to Biratnagar and Brt to Kathmandu



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Application



City Weather Map

City Pollution Map

Air Quality Modelling

Web based Visualization of Pollution
and Weather Proxies

Web based Visualization of
Pollution and Weather Proxies

Traffic Congestion Analysis

Bigdata Warehousing of Public
Remote Sensor Data

Routes from GPS Traces

Data Source for Further Researches

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

References



- “ K. Aberer, S. Sathe, D. Chakraborty, A. Martinoli, G. Barrenetxea, B. Faltings and L. Thiele, "OpenSense: open community driven sensing of environment," in *Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming*, 2010.
- “ T. Larson, S. B. Henderson and M. Brauer, "Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression," *Environmental science & technology*, vol. 43, pp. 4672-4678, 2009.
- “ T. Guo, K. Iwamura and M. Koga, "Towards high accuracy road maps generation from massive GPS traces data," in *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, 2007.
- “ S. Chakraborty, N. K. Nagwani and L. Dey, "Weather Forecasting using Incremental K-Means clustering," *arXiv preprint arXiv:1406.4756*, 2014.
- “ N. Cressie, "The origins of kriging," *Mathematical geology*, vol. 22, pp. 239-252, 1990.



Much obliged!



UNIVERSITY GRANTS COMMISSION
विश्वविद्यालय अनुदान आयोग

**For your support
with UGC Collaborative Research Grant**

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Much obliged!



Technical Consultant



Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

Much obliged!



Dr Nanda Bikram Adhikari

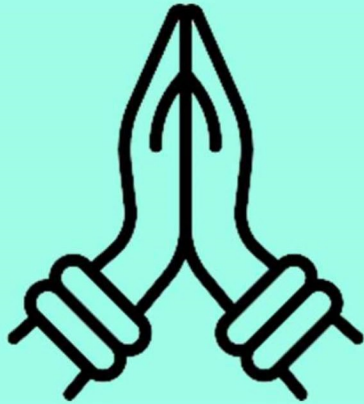
UGC Collaborative Research Grant

(Award, CRG-73/74-01Egg)

Near-Real Time Profiling of Fine Scale Environmental Proxies Using Mobile Sensors along Kathmandu Road Lines

ICQRIT 2019





Success is due to our stretching
to the challenges of life.
Failure comes when we
shrink from them.

John C. Maxwell



Thank You!

Alina Devkota	072 BCT 504	
Saloni Shikha	072 BCT 531	
Spandan Pyakurel	072 BCT 539	
Sushant Gautam	072 BCT 544	

