



Bicycle Usage Prediction

A DECISION TREE MODEL





ทำความเข้าใจ ธุรกิจ

ขั้นตอนที่ 1

ไม่ว่าทุกคนจะมียานพาหนะส่วนตัว โดยเฉพาะผู้ใช้ที่ต้องการจะเดินทางในระยะทางที่สั้น และไม่ยากจะพกจักรยานไปด้วยทุกที่ ธุรกิจการให้เช่าจักรยานจึงเข้ามาเพื่อตอบโจทย์ในส่วนๆ นี้ เราจึงอยากทำ model ที่สามารถทำนายอัตราจำนวนผู้ใช้บริการเช่าจักรยาน ในแต่ละสภาพอากาศและช่วงเวลา



ทำความเข้าใจข้อมูล

จากข้อมูลทั้งหมด 17 attributes

SEASON

ฤดูกาลที่บันทึกข้อมูลนั้น
(Spring, Summer, Fall,
Winter)

MONTH

เดือนที่บันทึกข้อมูล
(ตั้งแต่ 1-12)

HOURL

ชั่วโมงที่เก็บข้อมูล (0-23)

WORKINGDAY

วันนั้น เป็นวันที่คนออกไป
ทำงานหรือไม่? (0 = False, 1
= True)

WEATHERSIT

สภาพอากาศในวันๆ นั้น (1 -
แจ่มใส, 2 - เมฆเยาะ, 3 - ฝน
ตกอ่อน, 4 - ฝนตกหนัก)

TEMPERATURE

อุณหภูมิ ณ เวลานั้นๆ (หน่วย
เซลเซียส)

HUMIDITY

ความชื้น ณ เวลานั้นๆ

WIND SPEED

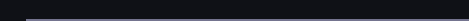
ความเร็วของลม ณ เวลาและ
จุดนั้นๆ (mph)

CLUSTER

จำนวนของผู้ใช้จากการคาด
การณ์ ซึ่งเป็น Label ที่ต้องการ
จะทำนาย (โดยแบ่งออกเป็น 5
กลุ่ม ได้แก่ Very High, High,
Average, Low, Very Low)

ID

Id ของข้อมูลนั้นๆ





ทำความเข้าใจ สะอาดข้อมูล

เราต้องเริ่มจากการจัดกลุ่มเป้าหมายของเรา ซึ่งนั่นก็คือจำนวนผู้ใช้รวมทั้งหมด เราจะแปลงจากเลขเพียวๆ ออกเป็นกลุ่มๆ ได้ 5 กลุ่ม นั่นก็คือ ผู้ใช้สูงมาก(Very High), มาก(High), ปานกลาง(Average), น้อย(Low) และ น้อยมาก(Very Low) แทนที่จะใช้เลขไปเลยจะช่วยให้โมเดลเราคาดคะเนได้แม่นยำกว่า และนำไปใช้จริงได้ง่ายกว่า

Generate

bicycle

Generate a new column below. You can type a formula or drag in columns from the left and functions from the right. Update the preview and - if all is correct - commit the result.

NAME: n_tkwl_count

Formula:

```
if([total_count]>=400,'very High',if([total_count]>=200,'High',if([total_count]>=100,'Average',if([total_count]>=50,'Low','Very Low'])))
```

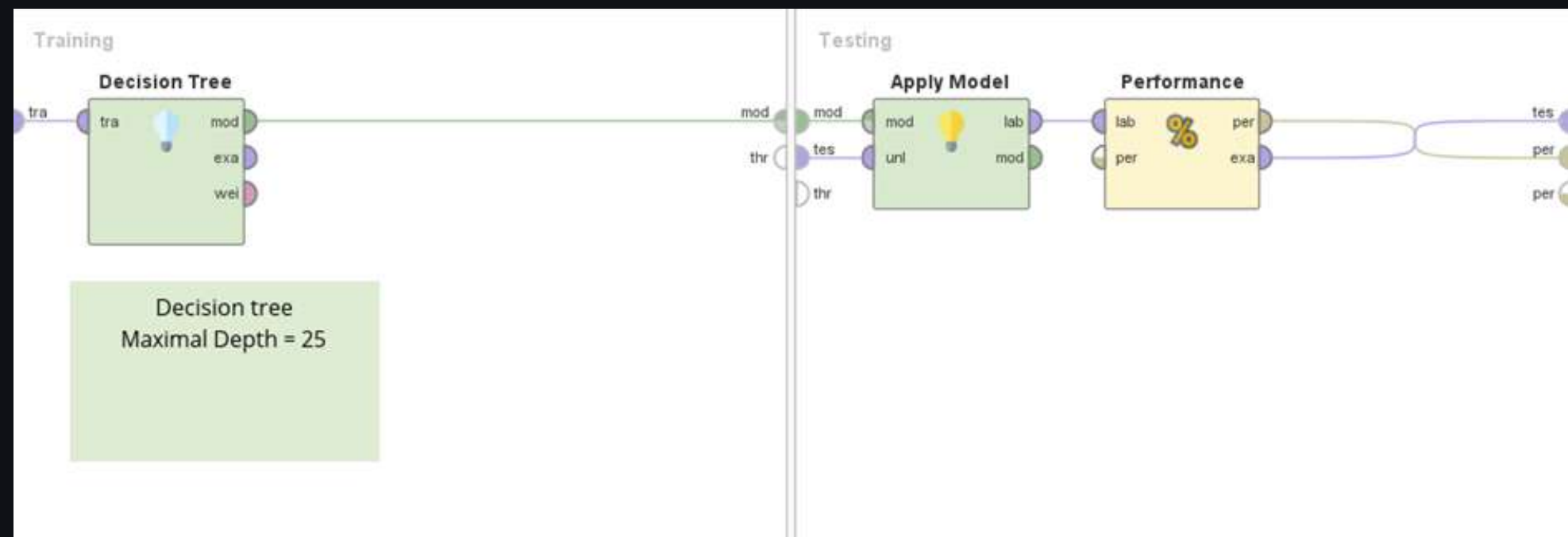
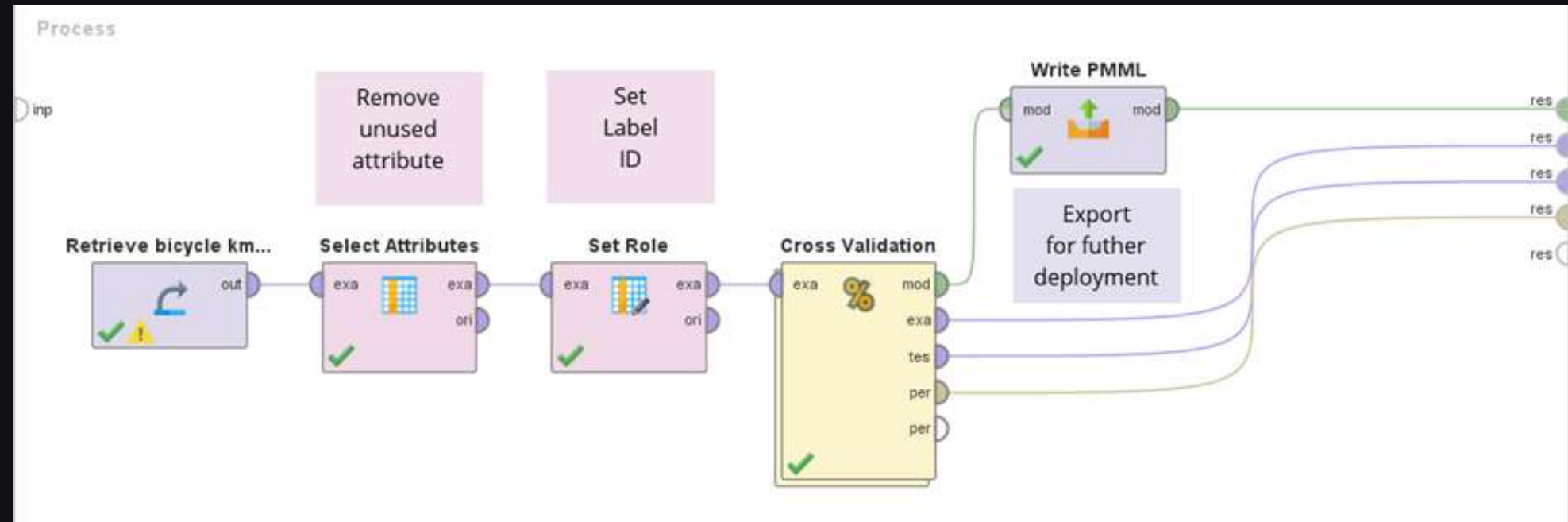
UPDATE PREVIEW - Update the preview to make sure that all calculations can be performed.

stamp	stamp	id	date	season	yr	month	hr	holiday
Normal	0.268	1	Jan 1, 2011	Spring	0	1	0	False
Normal	0.273	2	Jan 1, 2011	Spring	0	1	1	False
Normal	0.273	3	Jan 1, 2011	Spring	0	1	2	False
Normal	0.268	4	Jan 1, 2011	Spring	0	1	3	False
Normal	0.268	5	Jan 1, 2011	Spring	0	1	4	False
Normal	0.258	6	Jan 1, 2011	Spring	0	1	5	False
Normal	0.273	7	Jan 1, 2011	Spring	0	1	6	False
Normal	0.258	8	Jan 1, 2011	Spring	0	1	7	False

97,329 rows, 18 columns (3 nominal, 12 numerical, 1 date)

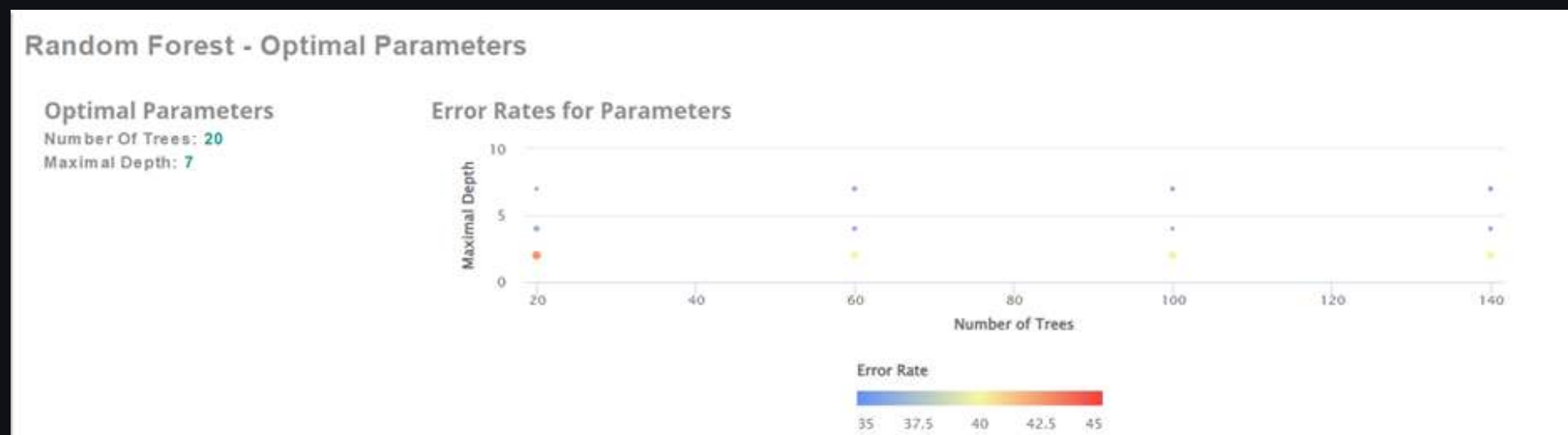


การสร้าง โมเดล

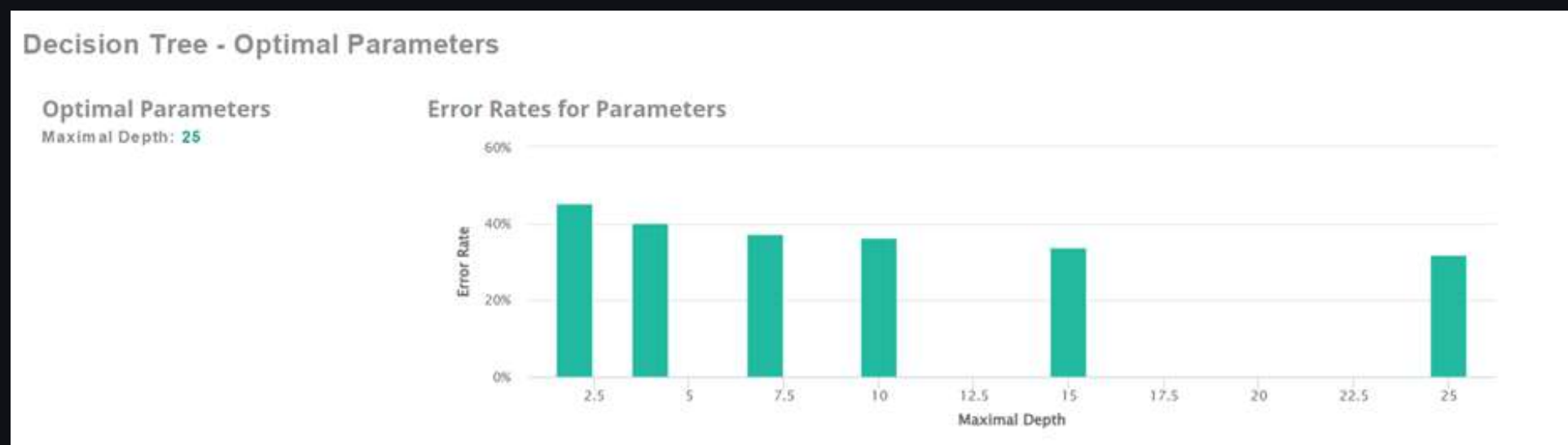




Best Parameter from Random forest – 39% error rate



Best Parameter from Decision tree – 31% error rate



ทำความเข้าใจ
ผลลัพธ์และ
ปรับปรุง



23 & 39



การนำไปใช้จริง

×

☐ Enable custom parameter?

Location (Input location name or coordinate)

มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

Format : Location Name

Refresh weather data

Predict

Temperature

29.51°C

Wind (Meters per hour)

3.09mph

Humidity

52%

Month : 2

Hour : 21

☒ Workingday?

Prediction

Average

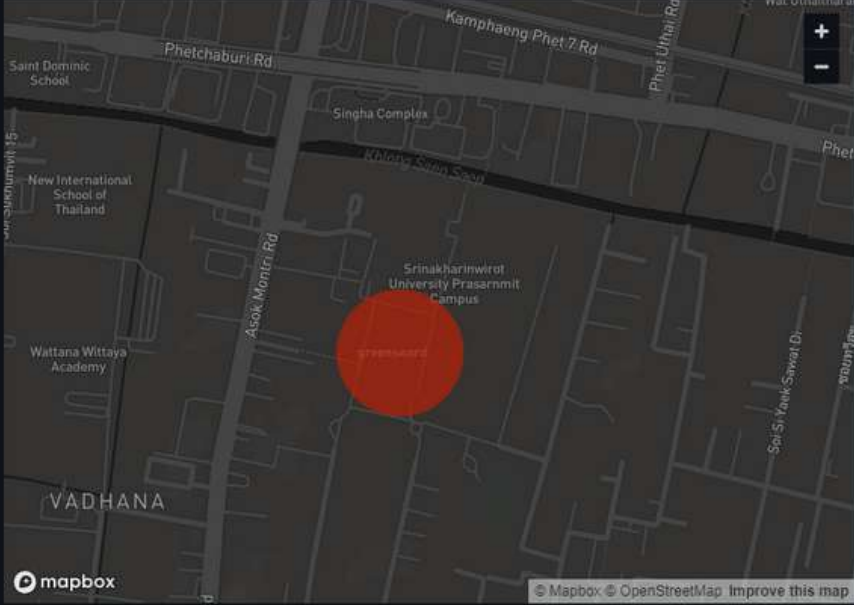
Welcome to bicycle usage prediction website!


Project by

> Jittaraboon Sapsintweelap No.23

> Tanapon Thongchamnong No.39

Data is pulled from [Bike Sharing Dataset](#)









Thank You!

WEB APP (Streamlit)



GitHub



PROJECT BY

JITTARABOON SAPSINTHAWHEELAP NO.23

TANAPON THONGCHAMNONG NO.39