# ANALYSIS OF FACTOR AFFECTING THE SOLID WASTE PRODUCTION AND MODEL PREDICTION USING MACHINE LEARNING AREA OF STUDY: BANGKOK

# การวิเคราะห์ปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอย และการทำนายด้วยแบบจำลองโดยใช้การเรียนรู้ของเครื่อง พื้นที่ศึกษา : กรุงเทพมหานคร

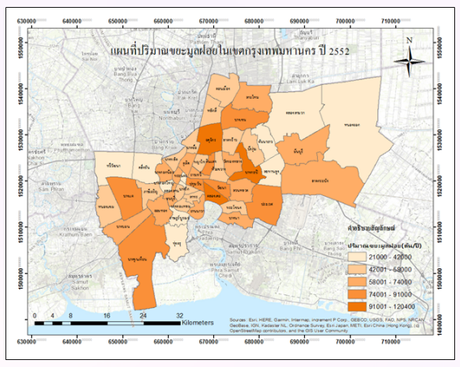
นางสาวสิประภา คะศรีทอง นางสาวอรอนงค์ ตระการจันทร์
สาขาวิชาภูมิศาสตร์และภูมิสารสนเทศ คณะมนุษยศาสตร์และสังคมศาสตร์

## บทคัดย่อ

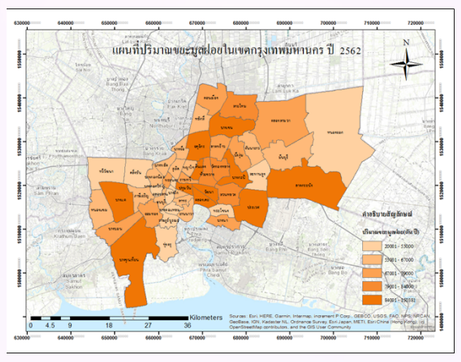
งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์หาปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยมากที่สุดในเขตกรุงเทพมหานคร โดยอาศัยระบบสารสนเทศภูมิศาสตร์ในการวิเคราะห์ข้อมูลปี พ.ศ. 2562 ปัจจัยที่นำมาวิเคราะห์มี 16 ปัจจัย ได้แก่ปริมาณขยะ จำนวนประชากร การใช้ประโยชน์ที่ดิน สถานีรถไฟฟ้า ที่ตั้งโรงงานอุตสาหกรรม ที่ตั้งศูนย์บริการเชื้อเพลิง ที่ตั้งสถานศึกษา ที่ตั้งสถานที่ราชการ พื้นที่ศูนย์บริการสุขภาพอนามัย ที่ตั้งวัด ที่ตั้งตลาด จำนวนครัวเรือน ที่ตั้งห้างสรรพสินค้า ที่ตั้งโรงแรม สนามกอล์ฟ และที่ตั้งชุมชน ทำการสร้างแผนที่แต่ละปัจจัยให้อยู่ในรูปแบบกริดขนาด 500x500 เมตร และแปลงข้อมูลปัจจัยให้ให้อยู่ในรูปแบบ Excel File เพื่อนำเข้าสู่กระบวนการถดถอยพหุคูณ ผลการศึกษาพบว่าปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยมากที่สุดประกอบด้วยปัจจัย จำนวนประชากร จำนวนครัวเรือน การใช้ประโยชน์ที่ดิน ที่ตั้งชุมชน ที่ตั้งโรงงานอุตสาหกรรม ที่ตั้งห้างสรรพสินค้า ที่ตั้งโรงแรม ที่ตั้งวัด สถานีรถไฟฟ้า ที่ตั้งสถานที่ราชการ พื้นที่ศูนย์บริการสุขภาพอนามัย โดยค่าสัมประสิทธิ์การตัดสินใจอยู่ที่ 77%จากนั้นนำปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยมากที่สุดมาวิเคราะห์แบบจำลองการเรียนรู้ของเครื่องในการคาดการณ์ปริมาณขยะมูลฝอยในเขตกรุงเทพมหานคร ผลการศึกษาพบว่า แบบจำลองซัพพอร์ทเวกเตอร์แมชชีน (SVM) , แบบจำลองดิซชั่นทรี(DT) , แบบจำลองนาอีฟเบย์(NB) และแบบจำลองแรนดอมฟอเรส(RF) ให้ค่าความถูกต้องที่ 83.4% 84.4% 95.5% และ 98.6% ตามลำดับ จากนั้นนำองค์ความรู้ของแบบจำลองในพื้นที่กรุงเทพมหานครไปตรวจสอบกับพื้นที่อื่น คือ พื้นที่จังหวัดนนทบุรีและพื้นที่จังหวัดชลบุรี ผลการศึกษาพบว่าองค์ความรู้ให้ค่าความถูกต้องที่ยอมรับได้ ผลการศึกษาครั้งนี้สามารถนำไปใช้วางแผนการจัดการปริมาณขยะเพื่อลดการเกิดปริมาณขยะที่อาจเพิ่มขึ้นรวมทั้งป้องกันการเกิดขยะที่ล้นเมืองได้ในอนาคต

## บทนำ

ระบบการรองรับการให้บริการการจัดการขยะมูลฝอยยังไม่เพียงพอในแหล่งท่องเที่ยว เนื่องจากไม่ได้มีการวางแผนและเตรียมความพร้อมเพื่อรองรับการจัดการขยะมูลฝอยควบคู่กับการพัฒนาแหล่งท่องเที่ยว รวมถึงมีการใช้สินค้าและบรรจุภัณฑ์ที่กำจัดยากอย่างฟุ่มเฟือย และเนื่องด้วยการบริหารจัดการได้ไม่เต็มประสิทธิภาพจึงเป็นเหตุให้ปริมาณขยะมีอัตราที่เพิ่มขึ้น (กองนโยบายและแผนงาน สำนักสิ่งแวดล้อม,2559) ดังนั้นทางคณะผู้จัดทำจึงเล็งเห็นความสำคัญของปัญหาขยะนี้ จึงทำการศึกษาเกี่ยวกับการวิเคราะห์ปัจจัยที่ส่งผลต่อการเปลี่ยนแปลงปริมาณขยะ โดยเลือกพื้นที่ศึกษากรุงเทพมหานครเป็นพื้นที่ศึกษา เนื่องจากมีปริมาณขยะมูลฝอยเกิดขึ้นต่อวันมากที่สุดในประเทศไทย จึงต้องมีการเตรียมการเพื่อรองรับสถานการณ์ขยะมูลฝอยที่จะเกิดขึ้นในกรุงเทพมหานคร เพื่อการวางแผนการจัดการขยะมูลฝอยให้เหมาะสม อีกทั้งยังสนับสนุนนโยบายและแผนงานของกรุงเทพมหานครด้านมาตรการการส่งเสริมบทบาทของทุกภาคส่วนทั้งภาครัฐ ภาคเอกชน (สำนักสิ่งแวดล้อม, 2563)




ที่มา: สำนักสิ่งแวดล้อม กรุงเทพมหานคร (2552)




ที่มา: สำนักสิ่งแวดล้อม กรุงเทพมหานคร (2562)

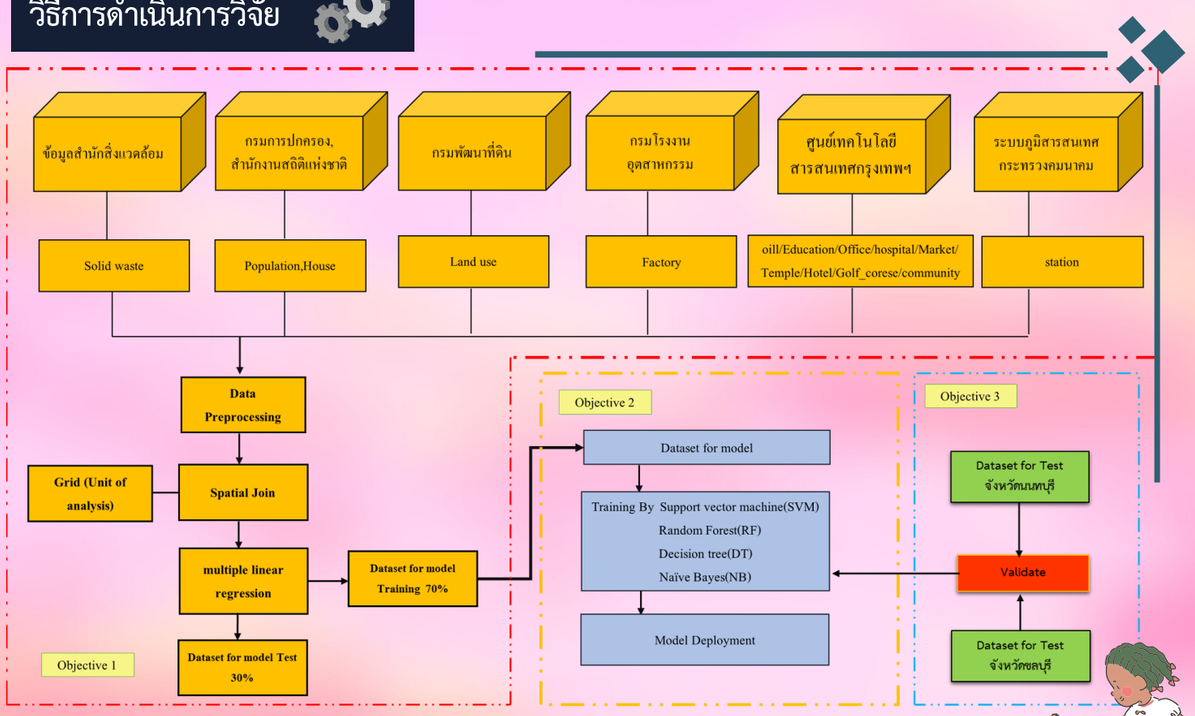
## วัตถุประสงค์ของการวิจัย

1.เพื่อศึกษาปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยในกรุงเทพมหานคร

2.เพื่อศึกษาแบบจำลองที่เหมาะสมในการคาดการณ์ปริมาณขยะมูลฝอยในกรุงเทพมหานคร

3.ตรวจสอบความถูกต้องของแบบจำลองที่ใช้ในการคาดการณ์ปริมาณขยะมูลฝอย

## วิธีการดำเนินการวิจัย




## ผลการวิจัย

**ผลการดำเนินการวิจัยตามวัตถุประสงค์ที่1 เพื่อศึกษาปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยในกรุงเทพมหานคร**

จากข้อมูลปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยทั้งหมด 16 ปัจจัย ดังรูปที่ 4 พบว่าปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยได้แก่ 11 ปัจจัย คือ จำนวนประชากร จำนวนครัวเรือน การใช้ประโยชน์ที่ดิน ที่ตั้งชุมชน ที่ตั้งโรงงานอุตสาหกรรม ห้างสรรพสินค้า โรงแรม ที่ตั้งโรงเรียน สถานีรถไฟฟ้า สถานที่ราชการ และพื้นที่ศูนย์บริการสุขภาพอนามัยโดยค่าสัมประสิทธิ์การตัดสินใจอยู่ที่ 77% จากการวิเคราะห์ปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยในเขตกรุงเทพมหานครโดยใช้แบบจำลอง Linear regression พบว่า มีค่าความเหมาะสมที่คำนวณค่า R-Square ได้ 0.77 ถือว่าอยู่ในเกณฑ์ที่ยอมรับได้

**ตารางที่ 1 ปัจจัยที่ส่งผลทำให้เกิดขยะมูลฝอยโดยแบบจำลอง Linear Regression**

| Factor | Estimate | Std. Error | T value | Pr(>ItI) |
|---|---|---|---|---|
| Incepter (ค่าคงที่) | -0.765088 | 0.073773 | -10.371 | $<2x10^{-16}$ |
| Population (ประชากร) | 0.200117 | 0.007893 | 25.354 | $<2x10^{-16}$ |
| House (ครัวเรือน) | 0.155891 | 0.005347 | 29.154 | $<2x10^{-16}$ |
| Land use(การใช้ประโยชน์ที่ดิน) | 0.079744 | 0.005356 | 14.889 | $<2x10^{-16}$ |
| Community (ชุมชน) | -0.021142 | 0.005360 | -3.945 | $8.06x10^{-5}$ |
| Factory (โรงงาน) | 0.059986 | 0.004975 | 12.056 | $<2x10^{-16}$ |
| Department (ห้างสรรพสินค้า) | 0.219310 | 0.054131 | 4.051 | $5.14x10^{-5}$ |
| Hotel (โรงแรม) | 0.271098 | 0.047132 | 5.752 | $9.15x10^{-9}$ |
| School (โรงเรียน) | -0.028837 | 0.009417 | -3.062 | 0.00220 |
| Station (สถานีรถไฟฟ้า) | 0.147415 | 0.054131 | 2.699 | 0.00697 |
| Office (สถานที่ราชการ) | 0.109060 | 0.016957 | 6.431 | $1.34x10^{-10}$ |
| Health Service Area (พื้นที่ศูนย์บริการสุขภาพอนามัย) | 0.642709 | 0.006618 | 97.113 | $<2x10^{-16}$ |

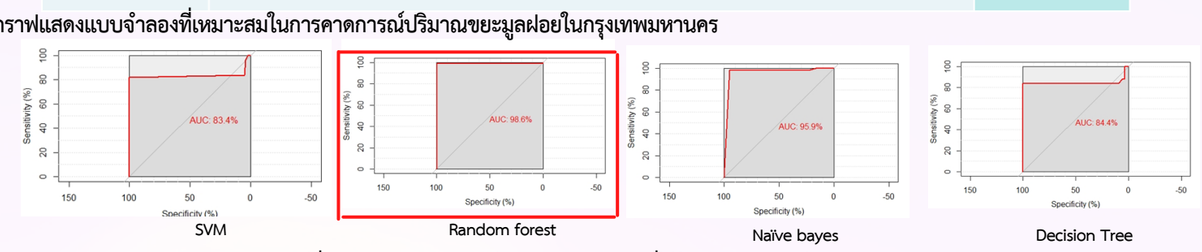
**ผลการดำเนินการวิจัยตามวัตถุประสงค์ที่ 2 เพื่อศึกษาแบบจำลองที่เหมาะสมในการคาดการณ์ปริมาณขยะมูลฝอยในกรุงเทพมหานคร**

จากข้อมูลปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยมากที่สุดได้แก่ 11 ปัจจัย คือ จำนวนประชากร จำนวนครัวเรือน การใช้ประโยชน์ที่ดิน ที่ตั้งชุมชน ที่ตั้งโรงงานอุตสาหกรรม ห้างสรรพสินค้า โรงแรม ที่ตั้งโรงเรียน สถานีรถไฟฟ้า สถานที่ราชการ และพื้นที่ศูนย์บริการสุขภาพอนามัยมาวิเคราะห์แบบจำลองการเรียนรู้ของเครื่องที่แม่นยำที่สุดในการคาดการณ์ปริมาณขยะมูลฝอยในกรุงเทพมหานคร ผลการศึกษาพบว่า แบบจำลองนาอีฟเบย์(NB) , แบบจำลองซัพพอร์ทเวกเตอร์แมชชีน (SVM) , แรนดอมฟอเรส(RF) และแบบจำลองดิซชั่นทรี(DT) ให้ค่าความถูกต้องที่ 95.5% 83.4% 98.6% และ 84.4% ตามลำดับ ดังตารางที่2

**ตารางที่ 2 แบบจำลองที่เหมาะสมในการคาดการณ์ปริมาณขยะมูลฝอยในกรุงเทพมหานคร**

| model | Algorithm | Accuracy |
|---|---|---|
| SVM | y~area+user+site+population+home+landuse+community+factory +department+hotel+temple+station+office+health service | 83.4 |
| Random forest | y~area+user+site+population+home+landuse+community+factory +department+hotel+temple+station+office+health service | 98.6 |
| Naïve bayes | y~area+user+site+population+home+landuse+community+factory +department+hotel+temple+station+office+health service | 95.9 |
| Decision Tree | y~area+user+site+population+home+landuse+community+factory +department+hotel+temple+station+office+health service | 84.4 |

**กราฟแสดงแบบจำลองที่เหมาะสมในการคาดการณ์ปริมาณขยะมูลฝอยในกรุงเทพมหานคร**




SVM Random forest Naïve bayes Decision Tree

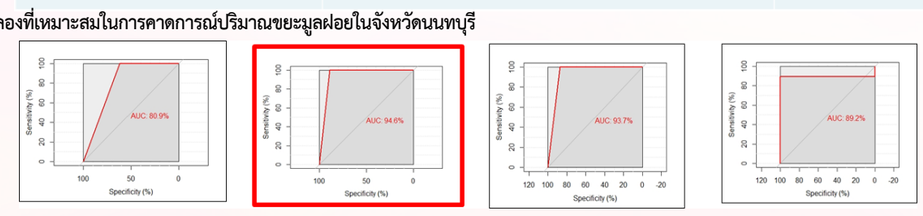
**ผลการดำเนินการวิจัยตามวัตถุประสงค์ที่ 3 ตรวจสอบความถูกต้องของแบบจำลองที่ใช้ในการคาดการณ์ปริมาณขยะมูลฝอย**

โดยนำชุดข้อมูลของพื้นที่จังหวัดนนทบุรี และเลือกปัจจัยที่มีความสัมพันธ์ต่อการเกิดปริมาณขยะมากที่สุด โดยคัดเลือกปัจจัยทั้งหมด 4 ปัจจัย ได้แก่ ประชากร ครัวเรือน ชุมชน และที่ตั้งสถานศึกษา มาเป็นตัวทดสอบ ผลการศึกษาพบว่า แบบจำลองซัพพอร์ทเวกเตอร์แมชชีน (SVM) , แบบจำลองดิซชั่นทรี(DT) , แบบจำลองนาอีฟเบย์(NB) และแบบจำลองแรนดอมฟอเรส(RF) ให้ค่าความถูกต้องที่ 80.9% 89.2% 93.6% และ 94.6% ตามลำดับ

**ตารางที่ 3 แบบจำลองที่เหมาะสมในการคาดการณ์ปริมาณขยะมูลฝอยในจังหวัดนนทบุรี**

| model | Algorithm | Accuracy |
|---|---|---|
| SVM | Y~Population+Home+com+school | 80.9 |
| Random forest | Y~Population+Home+com+school | 94.6 |
| Naïve bayes | Y~Population+Home+com+school | 93.7 |
| Decision Tree | Y~Population+Home+com+school | 89.2 |

**กราฟแสดงแบบจำลองที่เหมาะสมในการคาดการณ์ปริมาณขยะมูลฝอยในจังหวัดนนทบุรี**




SVM Random forest Naïve bayes Decision Tree

## สรุปผลและข้อเสนอแนะ

**อภิปรายผล**

การศึกษาปัจจัยที่ส่งผลต่อการเกิดขยะมูลฝอยโดยใช้กระบวนการวิเคราะห์ถดถอยเชิงพหุคูณ (Multiple Linear Regression) พบว่าปัจจัยที่ส่งผลมี 11 ปัจจัย ได้แก่ จำนวนประชากร จำนวนครัวเรือน การใช้ประโยชน์ที่ดิน ที่ตั้งชุมชน ที่ตั้งโรงงานอุตสาหกรรม ห้างสรรพสินค้า โรงแรม ที่สถานศึกษา สถานีรถไฟฟ้า สถานที่ราชการ พื้นที่ศูนย์บริการสุขภาพอนามัย ค่าสัมประสิทธิ์การตัดสินใจอยู่ที่ 77% แบบจำลองที่ให้ค่าความถูกต้องแม่นยำที่สุดในการทำนายปริมาณขยะมูลฝอยในเขตกรุงเทพมหานครคือ แบบจำลองแรนดอมฟอเรส (Random Forest) ให้ค่าความถูกต้องที่ 98.6% และได้นำไปตรวจสอบกับพื้นที่เขตในจังหวัดนนทบุรี และจังหวัดชลบุรีโดยผลการทดสอบจาก 4 ปัจจัยซึ่งให้ค่าความถูกต้องที่รับได้ เพราะปัจจัยที่ใช้ในการวิคราะห์มีความคล้ายคลึงกันกับพื้นที่เขตในกรุงเทพมหานคร

**ข้อเสนอแนะ**

1.ในการวิจัยครั้งนี้ เป็นการวิเคราะห์ปัจจัยที่ส่งผลต่อปริมาณขยะมูลฝอยโดยรวมทุกประเภท สำหรับผู้ที่สนใจพัฒนาต่อยอดงานวิจัยเรื่องนี้ควรทำการจำแนกปัจจัยที่ส่งผลต่อปริมาณขยะมูลฝอยของแต่ละประเภท และแยกปริมาณขยะแต่ละประเภท ใช้แบบสอบถามในการสำรวจแหล่งกำเนิดขยะมูลฝอยแต่ละกิจกรรมด้วยเพื่อให้การวิเคราะห์ปัจจัยที่ส่งผลต่อการขยะมูลฝอยมีความถูกต้องและใกล้เคียงความเป็นจริงมากที่สุด

2.หน่วยงานที่เกี่ยวข้องสามารถนำผลการศึกษาไปใช้ในการวางแผนจัดการและป้องกันการเพิ่มขึ้นของปริมาณขยะมูลฝอย

## เอกสารอ้างอิง

กระทรวงทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม . (2563, มกราคม 14) . สถานการณ์มลพิษของประเทศไทย ปี 2562 . สืบค้นจาก https://gnews.apps.go.th/

จตุรงค์ พะยอมแย้ม, กาญจนา นาคะภากร . (กันยายน 2554) . ระบบภูมิสารสนเทศเพื่อการคาดการณ์ปริมาณขยะจากการขยายตัวของเมือง . วารสารวิทยาศสตร์และเทคโนโลยี . 51-52

ยุทธ ไกยวกรรณ์ . (2555) . หลักการและการใช้การวิเคราะห์การถดถอยโลจิสติค . 1-12

สำนักสิ่งแวดล้อม . (2563) . แผนปฏิบัติการประจำปี พ.ศ.2563 . 6-9

Sonali Dubey . (2563) . Household Waste Management System Using IoT and Machine Learning