

การประมาณช่องสัญญาณโดยใช้วิธีการเรียนรู้ของเครื่อง
ในระบบไมโมขนานใหญ่



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมโทรคมนาคมและคอมพิวเตอร์
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2563

**CHANNEL ESTIMATION USING MACHINE
LEARNING FOR MASSIVE MIMO SYSTEMS**



Chittapon Keawin

**A Thesis Submitted in Partial Fulfillment of the Requirement for the
Degree of Master of Engineering in
Telecommunication and Computer Engineering
Suranaree University of Technology
Academic Year 2020**

การประเมินช่องสัญญาณโดยใช้วิธีการเรียนรู้ของเครื่อง
ในระบบโมเมนตัมใหญ่

มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์



(รศ. ดร.ปิยาภรณ์ มีสวัสดิ์)

ประธานกรรมการ



(รศ. ดร.พีระพงษ์ อุฑารสกุล)

กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)



(ผศ. ดร.อภิญา ตีค่ายคำ)

กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม)



(ผศ. ดร.ชิตพงษ์ เวชไชสงค์)

กรรมการ



(ผศ. ดร.คมเดช ภาพัฒน์บุรี)

กรรมการ



(รศ. ร.อ. ดร.กนต์ธร ชำนิประศาสน์)

รองอธิการบดีฝ่ายวิชาการและ

พัฒนาความเป็นสากล



(รศ. ดร.พรศิริ จงกล)

คณบดีสำนักวิชาวิศวกรรมศาสตร์

นายชิตกณ แก้วอินทร์ : การประมาณช่องสัญญาณโดยใช้วิธีการเรียนรู้ของเครื่องในระบบ
โมโฆฆานานใหญ่ (CHANNEL ESTIMATION USING MACHINE LEARNING FOR
MASSIVE MIMO SYSTEMS) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.
พีระพงษ์ อุฑารสกุล, 66 หน้า.

ในยุคปัจจุบันเทคโนโลยีมีการพัฒนาอย่างต่อเนื่อง จำนวนผู้ใช้งานเครือข่ายไร้สายเพิ่มขึ้น
เป็นจำนวนมาก จึงมีการใช้สายอากาศส่งมากกว่า 1 ต้น และสายอากาศรับมากกว่า 1 ต้น ภายใต้
ช่องสัญญาณวิทยุเดียวกัน โดยใช้หลักการเพิ่มจำนวนของเสาอากาศที่แยกกันรับและส่งออกจากกัน
เรียกว่าระบบ MIMO (Multiple Input Multiple Output) มารองรับผู้ใช้งานที่มากขึ้น แต่ก็ยังไม่
เพียงพอต่อความต้องการของผู้ใช้งานที่เพิ่มมากขึ้นในอนาคตที่จะรองรับระบบ 5G เพื่อรองรับ
เทคโนโลยีที่ทันสมัย เพิ่มระดับความเร็วในการส่งข้อมูล และลดข้อผิดพลาดที่อาจเกิดขึ้นในอนาคต
ทางผู้วิจัยเห็นความสำคัญของระบบ Massive MIMO คือระบบ MIMO ที่มีการเพิ่มปริมาณของ
สายอากาศมากขึ้น เช่น 32 หรือ 64 หรือมากกว่า

วิธีการประมาณช่องสัญญาณที่ใช้อย่างแพร่หลาย ได้แก่ เทคนิคการประมาณช่องสัญญาณ
Least Square (LS) มีการคำนวณ โดยไม่นำสัญญาณรบกวนมาพิจารณา และเทคนิคการประมาณ
ช่องสัญญาณ Minimum Mean Square Error (MMSE) มีการคำนวณ โดยนำสัญญาณรบกวนมา
พิจารณาร่วม อย่างไรก็ตามทั้งสองเทคนิคเป็นพื้นฐานที่ถูกใช้มานาน ความสามารถที่มีอาจจะยังไม่
เพียงพอต่อการประมาณช่องสัญญาณในระบบ Massive MIMO ดังนั้นจึงหาวิธีที่จะพัฒนาระบบให้
ลดความผิดพลาดและเพิ่มประสิทธิภาพให้ระบบดีขึ้น โดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine
Learning : ML) มาเปรียบเทียบเพื่อในการประมาณช่องสัญญาณ เพื่อลดความผิดพลาดที่เกิดขึ้น

เทคนิคการเรียนรู้ของเครื่อง (ML) มีบทบาทในการเพิ่มประสิทธิภาพงานในด้านต่าง ๆ เช่น
ทางการแพทย์ การขนส่ง และการสื่อสาร เพื่อมองหาอัลกอริทึมที่เหมาะสม ผู้วิจัยจึงหาอัลกอริทึม
ในกลุ่มของ Extreme Learning Machine (ELM) ที่ประกอบด้วย ELM, Regularized ELM (RELM)
และ Outlier Robust ELM (ORELM) ที่มีคุณสมบัติการวิเคราะห์ถดถอยทางข้อมูลสูง เพราะข้อมูล
จากช่องสัญญาณจำนวนมากเป็นค่าที่ไม่เท่ากันแต่มีความสัมพันธ์กันรวมถึงความสามารถในการ
เรียนรู้รวดเร็วและมีความซับซ้อนน้อย มาประยุกต์ใช้ในงานวิจัยร่วมกับระบบ Massive MIMO เพื่อ
ปรับปรุงประสิทธิภาพในระบบให้ดียิ่งขึ้น

จากนั้นนำผลที่ได้ทำการเปรียบเทียบกับการประมาณช่องสัญญาณกับเทคนิค LS และ MMSE ที่กล่าวมา เพื่อหาค่าความคลาดเคลื่อนเฉลี่ย (MSE) และอัตราความผิดพลาดบิต (BER) ในระบบ Massive MIMO



สาขาวิชา วิศวกรรมโทรคมนาคม
ปีการศึกษา 2563

ลายมือชื่อนักศึกษา โทชาล
ลายมือชื่ออาจารย์ที่ปรึกษา ดร. อนุสรณ์
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม อ. ปณิชา

CHITTAPON KEAWIN : CHANNEL ESTIMATION USING MACHINE
LEARNING FOR MASSIVE MIMO SYSTEMS. THESIS ADVISOR : ASSOC.
PROF. PEERAPONG UTHANSAKUL, Ph.D., 66 PP.

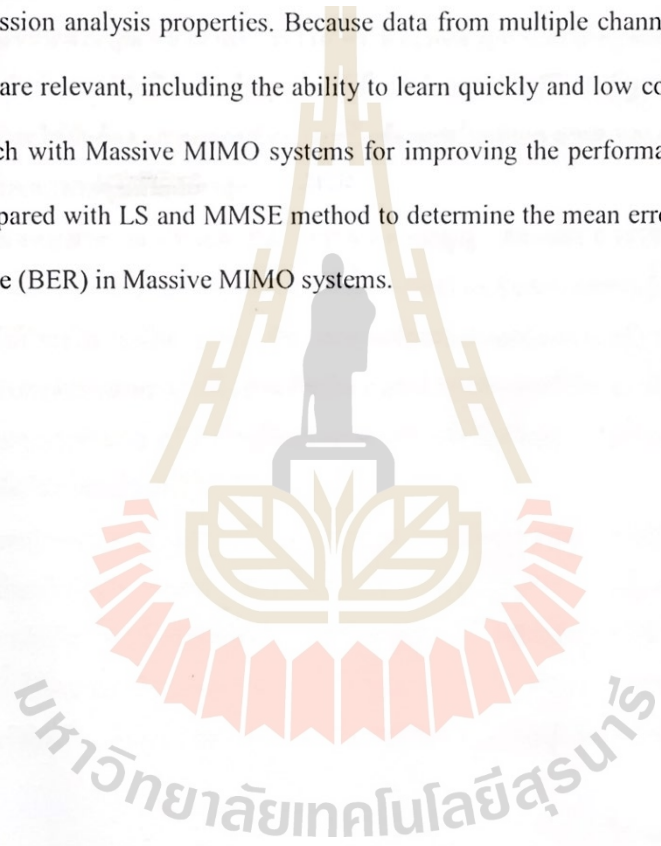
LS/ MMSE/ ELM/ RELM/ ORELM/ Channel Estimation/ Machine Learning

Nowadays, technology is constantly evolving. The number of wireless network users has increased dramatically. Therefore, more than one transmission antenna and more than one receiver antenna are used on the same radio channel. By using the principle of increasing the number of antennas that are separated to receive and transmit from each other, called MIMO (Multiple Input Multiple Output) systems to support more users. But it is still not enough to meet the needs of the growing users in the future to support the 5G system to support modern technology, increase the speed of data transmission and reduce errors that may occur in the future. The researchers see the importance of Massive MIMO systems as MIMO systems with increased antenna volume, such as 32 or 64 or more.

The most widely used channel estimation method, the Least Square (LS) channel estimation technique is calculated without noise into account and the Minimum Mean Square Error (MMSE) channel estimation technique is calculated by taking the noise into account. However, both techniques are fundamentals that have been used for a long time. The available capabilities may not be sufficient for channel estimation in a Massive MIMO system, therefore finding ways to improve the system to reduce errors and improve system performance. By using, the Machine Learning (ML) techniques to compare channel estimation for reducing error that occur.

compare channel estimation for reducing error that occur.

Machine learning (ML) techniques play a role in optimization, such as medical, transportation or communication to find the suitable algorithm. Therefore, the researchers looked for algorithms in the Extreme Learning Machine (ELM) group consisting of ELM, Regularized ELM (RELM), and Outlier Robust ELM (ORELM) with high regression analysis properties. Because data from multiple channels is not equal, but they are relevant, including the ability to learn quickly and low complexity using in research with Massive MIMO systems for improving the performance. The results are compared with LS and MMSE method to determine the mean error (MSE) and bit error rate (BER) in Massive MIMO systems.



School of Telecommunication Engineering Student's Signature

Academic year 2020 Advisor's Signature

Advisor's Asst. Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้มีจุดมุ่งหมายเพื่อศึกษาการประมาณช่องสัญญาณเพื่อใช้ในระบบสื่อสารและสำเร็จลงได้ด้วยดีเนื่องจากได้รับความกรุณาอย่างสูงจากทุกคนที่มีส่วนร่วมอุปการะอบรมสั่งสอนรวมถึงครอบครัวที่คอยให้กำลังใจเสมอมา

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.พีระพงษ์ อุฑารสกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ การดำเนินการวิจัยมีอาจสำเร็จลุล่วงไปได้หากปราศจากคำปรึกษาปัญหาในงานวิจัย และช่วยแนะนำและให้ข้อคิดจนสามารถปรับปรุงแก้ไขข้อบกพร่องต่างๆ ด้วยความเอาใจใส่เป็นอย่างดี จึงขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.อภิญา ดิถายคำ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ที่คอยให้คำปรึกษาในหลายๆ อย่างทั้งงานวิชาการหรือด้านเอกสารต่าง ๆ รวมถึงเป็นแรงผลักดันให้กำลังใจและเป็นแบบอย่างที่ควรยกย่อง จึงขอกราบขอบพระคุณเป็นอย่างยิ่ง

ขอกราบขอบพระคุณคณาจารย์ในสาขาวิชาวิศวกรรมโทรคมนาคมทั้งใน มหาวิทยาลัยและนอกมหาวิทยาลัยทุก ๆ ท่านที่ช่วยให้ความรู้ในด้านวิชาการ และให้คำแนะนำในการดำเนินงานต่างๆ รวมถึงให้ข้อคิดในการดำเนินชีวิต

ขอกราบขอบพระคุณบิดา มารดา ที่ให้การอุปการะอบรมเลี้ยงดูตลอดจนส่งเสริมการศึกษา และให้กำลังใจเป็นอย่างดี อีกทั้งขอขอบคุณพี่ เพื่อน และน้องทุกคนที่มีส่วนร่วมให้การสนับสนุน และให้ความช่วยเหลือในทุก ๆ เรื่องและคอยให้กำลังใจจนงานวิจัยนี้สำเร็จลุล่วงไปได้ด้วยดี

สุดท้ายนี้ผู้วิจัยหวังว่าวิทยานิพนธ์ฉบับนี้จะมีส่วนในการแก้ไขปัญหาและพัฒนา หรือมีประโยชน์ต่องานวิจัยที่เกี่ยวข้องไม่มากนักน้อย หากมีข้อบกพร่องประการใดผู้วิจัยก็ขออภัยไว้ ณ ที่นี้

ชิตภณ แก้วอินทร์

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ค
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญรูป.....	ฉ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตการดำเนินงาน.....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.5 วิธีดำเนินการวิจัย.....	3
1.6 ส่วนประกอบของวิทยานิพนธ์.....	3
1.7 สรุป.....	4
2 ปรัชญ์นั้วรรณกรรมและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 กล่าวนำ.....	5
2.2 MIMO และการมาของ Massive MIMO.....	5
2.3 การประมาณช่องสัญญาณ (Channel Estimation).....	7
2.4 การเรียนรู้ของเครื่อง (Machine Learning ; ML).....	12
2.5 สิ่งทีว้วิเคราะห์เปรียบเทียบ.....	19
2.6 ปรัชญ์นั้วรรณกรรมและงานวิจัย.....	27
2.7 สรุป.....	30
3 วิธีการดำเนินงานวิจัย.....	31
3.1 กล่าวนำ.....	31

สารบัญ (ต่อ)

หน้า

3.2	การจำลองระบบ.....	31
3.3	เทคนิคการเรียนรู้ของเครื่อง Extreme Learning Machine	33
3.4	เทคนิคการเรียนรู้ของเครื่อง Regularized Extreme Learning Machine.....	34
3.5	เทคนิคการเรียนรู้ของเครื่อง Outliner-Robust Extreme Learning Machine	36
3.6	การนำเทคนิคการเรียนรู้ของเครื่องมาประยุกต์หาการประมาณช่องสัญญาณ	38
3.7	การนำเทคนิคการประมาณช่องสัญญาณ LS และ MMSE มาใช้งาน	40
3.8	วิธีการเทียบประสิทธิภาพแบบ MSE	41
3.9	วิธีการเทียบประสิทธิภาพแบบ BER.....	41
3.10	วิธีการเทียบประสิทธิภาพแบบการหาความจุช่องสัญญาณ	42
3.11	สรุป.....	43
4	ผลการดำเนินงาน	44
4.1	กล่าวนำ.....	44
4.2	การทดสอบหาค่าความคลาดเคลื่อนของอัลกอริทึม.....	44
4.3	การทดสอบหาอัตราบิดผิดพลาดในการส่งข้อมูล	49
4.4	การทดสอบหาความจุช่องสัญญาณ.....	50
4.5	ผลการทดสอบการประมวลผลทางเวลา (Processing Times)	51
4.6	สรุปผลการดำเนินงาน	55
5	สรุปเนื้อหาวิทยานิพนธ์และข้อเสนอแนะในการวิจัย.....	56
5.1	สรุปเนื้อหาวิทยานิพนธ์	56
5.2	ข้อเสนอแนะและแนวทางการพัฒนาในอนาคต.....	57
	รายการอ้างอิง	58
	ภาคผนวก ก.....	60
	ประวัติผู้เขียน.....	66

สารบัญตาราง

ตารางที่	หน้า
2.1 การเปรียบเทียบประสิทธิภาพวิธีในการจำแนกและทำข้อมูล.....	17
4.1 ทดสอบระยะเวลาในการคำนวณค่าความคลาดเคลื่อน (MSE).....	52
4.2 ทดสอบระยะเวลาในการคำนวณอัตราความผิดพลาดข้อมูล (BER).....	52



สารบัญรูป

รูปที่	หน้า
2.1	รูปแบบการแพร่กระจายคลื่น โดยสายอากาศ Massive MIMO6
2.2	Spectral efficiency vs. Energy efficiency in SISO, MISO and Massive MIMO systems with different processing methods; source: [4]7
2.3	โครงสร้างระบบการสื่อสาร7
2.4	บล็อกไดอะแกรมของการประมาณช่องสัญญาณแบบ MMSE.....10
2.5	ประเภทของ Supervised Learning13
2.6	ตัวอย่างกราฟระหว่างอุบัติเหตและจำนวนประชากร14
2.7	การจำแนกข้อมูล (a).....15
2.8	การเรียนรู้จากข้อมูลเพื่อสร้างตัวจำแนกข้อมูล (b)15
2.9	การแบ่งกลุ่มของการจับกลุ่มข้อมูล18
2.10	การแสดงผลของการจัดข้อมูลแล้วนำไปแยกประเภท class [Figure credit: Sebastian Raschka, Python Machine Learning].....19
2.11	กลุ่มเครื่องมือวัด Classification and Regression Loss. (www.heartbeat.fritz.ai)20
2.12	กราฟของ MSE Loss (Y-axis) เทียบกับ Predictions (X-axis).....21
2.13	กราฟของ MAE Loss (Y-axis) เทียบกับ Predictions (X-axis)22
2.14	กราฟของ Huber Loss (Y-axis) เทียบกับ Predictions (X-axis).....23
2.15	กราฟ Log-cosh Loss (Y-axis) เทียบกับ Predictions (X-axis).....24
2.16	กราฟ Log-cosh Loss (Y-axis) เทียบกับ Predictions (X-axis).....25
2.17	อัตราบิดผิดพลาดสำหรับ BPSK , QPSK , 8-PSK และ 16-PSK โดยใช้ช่องสัญญาณ AWGN26
2.18	อัตราบิดผิดพลาดสำหรับ BPSK และ BPSK ที่แตกต่างกัน โดยใช้ช่องสัญญาณ AWGN26
2.19	ประสิทธิภาพ MSE โดยเทคนิค LS,LMMSE และ DL เทียบกับ SNR28
2.20	ประสิทธิภาพ BER โดยเทคนิค LS,LMMSE, DNN-1 และ DNN-2 เทียบ SNR29
2.21	ประสิทธิภาพ MSE โดยเทคนิค LS,LMMSE, DNN-1 และ DNN-2 เทียบ SNR29

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.1	โครงสร้างระบบประสาทของ Extreme Learning Machine32
3.2	โครงสร้างการทำงานในการเตรียมข้อมูล38
3.3	โครงสร้างการระบุข้อมูลที่ใช้ในกระบวนการเรียนรู้ของเครื่อง39
3.4	ลักษณะการส่งข้อมูลเข้ารหัสและถอดรหัสแบบง่าย41
4.1	ผลทดสอบ MSE ในระบบสายอากาศส่ง 16 ต้น และสายอากาศรับ 16 ต้น.....45
4.2	ผลทดสอบ MSE ในระบบสายอากาศส่ง 64 ต้น และสายอากาศรับ 64 ต้น.....45
4.3	ผลทดสอบ MSE ในระบบสายอากาศส่ง 128 ต้น และสายอากาศรับ 128 ต้น.....46
4.4	การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณ แบบ ELM, LS และ MMSE47
4.5	การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณ แบบ LS, MMSE, ELM และ RELM48
4.6	การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณ แบบ LS, MMSE, ELM, RELM และ ORELM49
4.7	การทดสอบหา BER เทียบกับ SNR.....50
4.8	ความจุช่องสัญญาณทั้ง 5 เทคนิคเทียบกับอัตราส่วนสัญญาณต่อสัญญาณรบกวน51
4.9	ระยะเวลาในการคำนวณทางเวลา ELM และ RELM.....53
4.10	ระยะเวลาในการคำนวณทางเวลา ELM, RELM และ ORELM.....54
4.11	ผลทดสอบผลลัพธ์ของการประเมินระหว่างตัวแปรและเวลาในช่วง Training Data55
4.12	ผลทดสอบผลลัพธ์ของการประเมินระหว่างตัวแปรและเวลาในช่วง Testing Data55

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันระบบการสื่อสารไร้สายมีความสำคัญ เนื่องจากมีผู้ใช้งานจำนวนมากในระบบต่าง ๆ เช่น อินเทอร์เน็ต โทรศัพท์เคลื่อนที่ การประชุมออนไลน์ หรือการสื่อสารอื่น ๆ ที่ต้องใช้เทคโนโลยีไร้สายเข้ามาเกี่ยวข้อง การเข้ามาของเทคโนโลยี 5G ทำให้เกิดการแข่งขันระหว่างผู้ให้บริการโครงข่ายต้องพัฒนาระบบให้มีประสิทธิภาพมากขึ้น เพื่อรองรับการใช้งานของคนจำนวนมากขึ้น จากเดิมที่ใช้สายอากาศเพียงต้นเดียวในการสื่อสารรับ-ส่งข้อมูล ที่เรียกว่าระบบ SISO (Single Input Single Output) ต่อมาเทคโนโลยีได้พัฒนามากขึ้นมีการส่งข้อมูลเพิ่มขึ้นอาจใช้สายอากาศส่งมากกว่า 1 ต้น และสายอากาศรับต้นเดียว เรียกว่าระบบ MISO (Multiple Input Single Output) หรือ สายอากาศรับมากกว่า 1 ต้น สายอากาศส่งแค่ต้นเดียว เรียกว่าระบบ SIMO (Single Input Multiple Output) และในยุคปัจจุบันเทคโนโลยีมีการพัฒนาอย่างต่อเนื่อง จำนวนผู้ใช้งานเครือข่ายไร้สายเพิ่มขึ้น จึงมีการใช้สายอากาศส่งมากกว่า 1 ต้น และสายอากาศรับมากกว่า 1 ต้น ภายใต้ช่องสัญญาณวิทยุเดียวกัน โดยใช้หลักการเพิ่มจำนวนของเสาอากาศที่แยกกันรับและส่งออกจากกัน เรียกว่าระบบ MIMO (Multiple Input Multiple Output) มารองรับผู้ใช้งานเพิ่มมากขึ้น และในปัจจุบันก็ยังไม่เพียงพอต่อความต้องการให้รองรับระบบในอนาคตที่ใช้เทคโนโลยี 5G ให้ได้ความเร็วที่เพิ่มขึ้นตามผู้ใช้งาน เพื่อเพิ่มระดับความเร็วในการส่งข้อมูล และลดข้อผิดพลาดที่อาจเกิดขึ้นในอนาคต ทางผู้วิจัยเห็นความสำคัญของระบบ Massive MIMO นำเอาระบบ MIMO ที่มีการเพิ่มปริมาณของสายอากาศมากขึ้น เช่น 8 ต้น 32 หรือ 64 หรือมากกว่า ในงานวิจัยนี้จะจงใจไปที่สายอากาศส่ง 128 ต้น และสายอากาศรับ 128 ต้น เพื่อรองรับความต้องการ การทำงานโครงข่ายขนาดใหญ่ในอนาคต เนื่องจากมีสายอากาศภาครับและสายอากาศภาคส่งจำนวนมาก ช่วยให้อัตราการส่งข้อมูลที่สูงแก่ผู้ใช้งานเป็นจำนวนมาก เพื่อเพิ่มขีดความสามารถ และรองรับบริการมัลติมีเดียระดับ real time โดยไม่จำเป็นต้องใช้คลื่นความถี่เพิ่มเติม ส่วนสำคัญที่ใช้พิจารณาการทำงานของระบบมีหลายปัจจัย เช่น การประมาณช่องสัญญาณเพื่อเพิ่มประสิทธิภาพในการส่งข้อมูลให้ดีขึ้น และลดข้อผิดพลาดของข้อมูลให้ลดลง วิธีการประมาณช่องสัญญาณที่ใช้อย่างแพร่หลาย ได้แก่ เทคนิคการประมาณช่องสัญญาณ Least Square (LS) มีการคำนวณ โดยไม่นำสัญญาณรบกวนมาพิจารณา และเทคนิคการประมาณช่องสัญญาณ Minimum Mean Square Error (MMSE) มีการ

คำนวณโดยนำสัญญาณรบกวนมาพิจารณาร่วม อย่างไรก็ตามทั้งสองเทคนิคเป็นพื้นฐานที่ ถูกใช้งาน ความสามารถที่มีอาจจะยังไม่เพียงพอต่อการประมาณช่องสัญญาณในระบบ Massive MIMO ดังนั้นทางผู้วิจัยจึงหาวิธีที่จะพัฒนาระบบให้ลดความผิดพลาดและเพิ่มประสิทธิภาพให้ ระบบดีขึ้น ทางผู้วิจัยจึงใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning : ML) เข้ามาพิจารณา เปรียบเทียบเพื่อใช้เป็นตัวเลือกใหม่ในการประมาณช่องสัญญาณ เพื่อลดความผิดพลาดที่เกิดขึ้นใน อนาคต

ปัจจุบันการเทคนิคการเรียนรู้ของเครื่อง (ML) เริ่มมีบทบาทในการเพิ่มประสิทธิภาพงาน ในด้านต่าง ๆ ไม่ว่าจะเป็นทางการแพทย์ การขนส่ง มีการศึกษาเทคนิคการเรียนรู้ของเครื่อง ได้แก่ Algorithm ที่ได้รับความนิยมสูง คือ Deep Learning ซึ่งถูกออกแบบมาให้ใช้งานและประยุกต์ได้ หลายลักษณะงาน อย่างไรก็ตาม ในการทำงานจริง Data Scientist จำเป็นต้องออกแบบตัวแปรต่าง ๆ ทั้งในตัวของ Deep Learning เอง และต้องหา Algorithm อื่น ๆ มาเป็นคู่เปรียบเทียบ เพื่อมองหา Algorithm ที่เหมาะสมที่สุด ทำให้แนวคิดนี้มีความซับซ้อน ดังนั้นผู้วิจัยจึงหา Algorithm ในกลุ่ม ของ Extreme Learning Machine (ELM) ที่ประกอบด้วย ELM, Regularized ELM (RELM) และ Outlier Robust ELM (ORELM) ที่มีคุณสมบัติการวิเคราะห์ที่ถดถอยสูงเพราะข้อมูลจากช่องสัญญาณ จำนวนมากเป็นค่าที่ไม่เท่ากันแต่มีความสัมพันธ์กันรวมถึงความสามารถในการเรียนรู้รวดเร็วและมี ประสิทธิภาพที่ไม่ซับซ้อน มาประยุกต์ใช้ในงานวิจัยร่วมกับระบบ Massive MIMO เพื่อปรับปรุง ประสิทธิภาพในระบบให้ดียิ่งขึ้น

จากนั้นนำผลที่ได้ทำการเปรียบเทียบกับการประมาณช่องสัญญาณทั้งสองที่กล่าวมา เพื่อหา ค่าความคลาดเคลื่อนเฉลี่ย (MSE) อัตราความผิดพลาดบิต (BER) และความจุช่องสัญญาณ(Channel Capacity) ในระบบ Massive MIMO รวมถึงมีการเปรียบเทียบกระบวนการทำงานทางเวลาของวิธี ELM, RELM และ ORELM โดยผลที่เกิดขึ้น เพื่อเป็นวิธีในการเลือกใช้งานแต่ละวิธีในอนาคต

1.2 วัตถุประสงค์ของการวิจัย

- 1.2.1 ประยุกต์ใช้การประมาณช่องสัญญาณ โดยใช้เทคนิคการเรียนรู้ของเครื่อง
- 1.2.2 ทดสอบ Algorithm ที่มีความสามารถในการประมาณช่องสัญญาณ
- 1.2.3 เปรียบเทียบประสิทธิภาพ Algorithm ที่มีความสามารถในการประมาณ ช่องสัญญาณทั่วไปและเทคนิคการประมาณช่องสัญญาณเรียนรู้ของเครื่อง

1.3 ขอบเขตการดำเนินงาน

- 1.3.1 ศึกษาเทคนิคการประมาณช่องสัญญาณในระบบ Massive MIMO

1.3.2 ประยุกต์ใช้การเรียนรู้ของเครื่องมากกว่า 1 ชนิด

1.3.3 เปรียบเทียบประสิทธิภาพที่ใช้ในการประมาณช่องสัญญาณ ด้วยการประยุกต์ใช้การเรียนรู้ของเครื่อง

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 ผลการเปรียบเทียบประสิทธิภาพที่ใช้ในการประมาณช่องสัญญาณ ด้วยการประยุกต์ใช้การเรียนรู้ของเครื่อง

1.4.2 ได้ศึกษาเทคนิคการเรียนรู้ของเครื่อง เพื่อเป็นต้นแบบในการพัฒนาต่อไป

1.5 วิธีดำเนินการวิจัย

1.5.1 สํารวจปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้องกับวิทยานิพนธ์

1.5.2 ศึกษาเทคนิคการประมาณช่องสัญญาณในระบบสื่อสารไร้สายระบบ Massive MIMO

1.5.3 ศึกษาการใช้ระบบการเรียนรู้ของเครื่อง (Machine Learning)

1.5.4 ประยุกต์ใช้การเรียนรู้ของเครื่องกับการประมาณช่องสัญญาณ

1.5.5 เปรียบเทียบประสิทธิภาพที่ใช้ในการประมาณช่องสัญญาณ ด้วยการประยุกต์ใช้การเรียนรู้ของเครื่อง

1.6 ส่วนประกอบของวิทยานิพนธ์

สำหรับเนื้อหาของวิทยานิพนธ์ฉบับนี้ นำเสนอการศึกษาค้นคว้า รวบรวมข้อมูล ออกแบบระบบ ทดสอบระบบ วิเคราะห์และสรุปผล โดยประกอบไปด้วยเนื้อหาทั้งหมด 5 บท

บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ของการวิจัย ขอบเขตการวิจัย ประโยชน์ที่คาดว่าจะได้รับ วิธีดำเนินการวิจัยของงานวิจัยนี้

บทที่ 2 กล่าวถึงแนวคิด ทฤษฎี เอกสารและงานวิจัยที่เกี่ยวข้องกับระบบเครือข่ายไร้สาย สำหรับการประมาณช่องสัญญาณ รวมไปถึงทฤษฎีของการเรียนรู้ของเครื่อง (Machine Learning)

บทที่ 3 กล่าวถึงวิธีดำเนินการออกแบบระบบการประมาณช่องสัญญาณในระบบโมโนมัลติแอนเทนนาขนาดใหญ่ (Massive MIMO) และใช้การประยุกต์ของวิธีการเรียนรู้ของเครื่องมาเพิ่มประสิทธิภาพ เช่นอัตราข้อผิดพลาดข้อมูลและความคลาดเคลื่อนของอัลกอริทึมที่ใช้ในการประมาณช่องสัญญาณ

บทที่ 4 กล่าวถึงผลการทดสอบการประมาณช่องสัญญาณที่ใช้เทคนิคแบบดั้งเดิมและการนำเอาเทคนิคการเรียนรู้ของเครื่องมาประยุกต์ใช้งาน เพื่อแสดงให้เห็นว่าประสิทธิภาพของการ

เรียนรู้ของเครื่องนั้นมีประสิทธิภาพและสามารถนำมาใช้ในการประมาณช่องสัญญาณในระบบโมโฆนนานใหญ่

บทที่ 5 กล่าวถึงการสรุปผลของการวิจัย ข้อเสนอแนะและแนวทางการพัฒนาต่อไปในอนาคต

1.7 สรุป

ความเป็นมาและปัญหาที่ทำให้เกิดงานวิจัยที่ได้นำเสนอในบทที่ 1 นี้ เป็นงานวิจัยที่เกี่ยวข้องเครือข่ายไร้สายในระบบโมโฆนนานใหญ่และการประมาณช่องสัญญาณที่ใช้กันมาอย่างยาวนาน ซึ่งการดำเนินงานวิจัยวิทยานิพนธ์จะทำการศึกษาระบบโมโฆนนานใหญ่และหาวิธีนำเอาการเรียนรู้ของเครื่องมาประยุกต์ใช้กับการประมาณ เพื่อให้สามารถลดข้อผิดพลาดในการรับส่งข้อมูลที่มีจำนวนมากและเป็นการพิสูจน์ประสิทธิภาพของการเรียนรู้ของเครื่องว่าเหมาะสมกับการประมาณช่องสัญญาณอีกด้วย



บทที่ 2

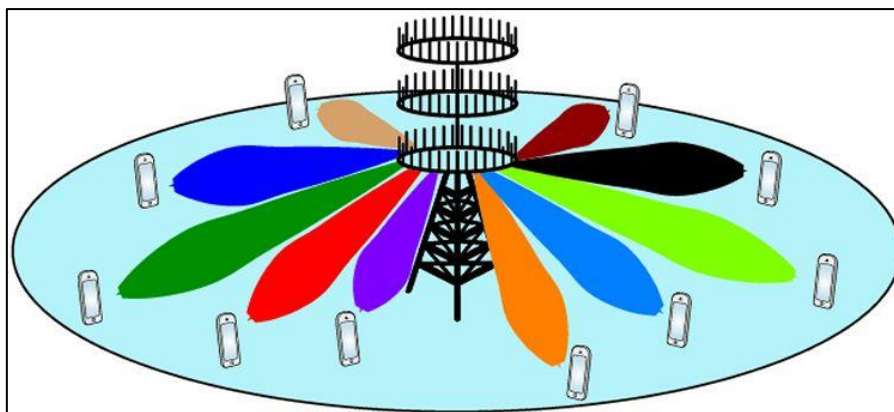
ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 กล่าวนำ

เทคโนโลยีของการสื่อสารในปี พ.ศ 2564 ได้ถูกพัฒนามาอย่างต่อเนื่อง เช่น การรับส่งข้อมูลและรับข้อมูลโดยใช้สายอากาศรับและส่งมากกว่า 1 ต้น (Multiple Input Multiple Output : MIMO) ซึ่งเป็นที่นิยมใช้งานอย่างกว้างขวางและได้รับความสนใจจากนักวิจัยทั่วโลก เนื่องจากความสามารถในการลดอัตราข้อผิดพลาดการรับส่งข้อมูลและการเพิ่มอัตราเร็วในการรับส่งข้อมูล ในปัจจุบันได้มีการพัฒนาอย่างต่อเนื่อง แต่ยังไม่เพียงพอต่อผู้ใช้งานในปัจจุบันที่เพิ่มมากขึ้น ทางผู้วิจัยจึงได้ศึกษาระบบ เพิ่มจำนวนสายอากาศส่งและรับ มากกว่า 8 ต้น เรียกว่าระบบโมฆอนานใหญ่ (Massive MIMO) โดยมีการศึกษางานวิจัยที่เกี่ยวข้องและได้นำเสนอในบทนี้ เพื่อแก้ปัญหาข้อจำกัดบางอย่างของระบบการรับส่งข้อมูล มีการอธิบายหลักการทำงานของ การประมาณช่องสัญญาณ (Channel Estimation) ในแต่ละเทคนิค เช่น เทคนิคการประมาณช่องสัญญาณ Least Square Estimation (LS) เทคนิคการประมาณช่องสัญญาณ Minimum Mean-Square Error Channel Estimation (MMSE) และเทคนิคการเรียนรู้ของเครื่อง (Machine Learning ; ML) รวมถึงการอธิบายพารามิเตอร์ที่ใช้ในการเปรียบเทียบเพื่อนำมาวิเคราะห์ผลที่เกิดขึ้น

2.2 MIMO และการมาของ Massive MIMO

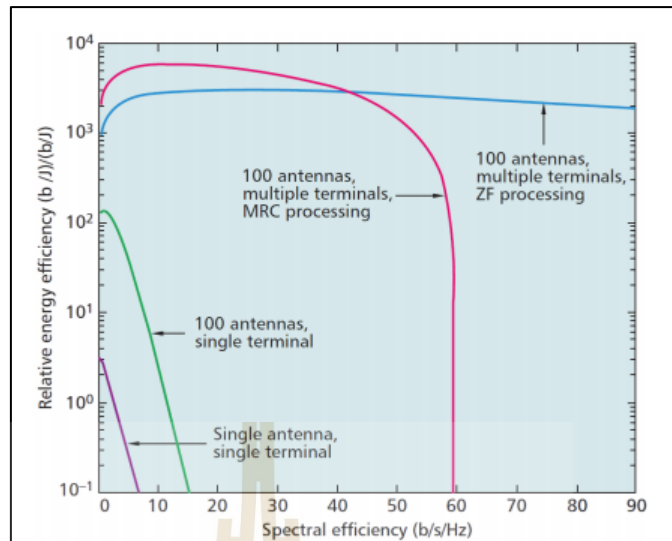
ในระบบสื่อสารทั่วไปนั้นส่วนแล้วเกิดมาจากความต้องการของผู้คนทั่วโลกที่ถูกพัฒนามาอย่างต่อเนื่อง โดย MIMO มีการใช้สายอากาศมากกว่า 1 ต้น ทั้งภาคส่งและภาครับ ซึ่งต่างจากเทคโนโลยีในยุคก่อน ๆ ที่ใช้สายอากาศทางภาคส่งหลายต้น ส่วนภาครับมีต้นเดียว หรือทางภาครับหลายต้นส่วนภาคส่งมีต้นเดียว แต่ก็ยังไม่เพียงพอต่อความต้องการของผู้ใช้งาน เพราะอาจจะเกิด Multipath fading ของสัญญาณทำให้การรับข้อมูลไม่ถูกต้องและล่าช้า



รูปที่ 2.1 รูปแบบการแพร่กระจายคลื่นโดยสายอากาศ Massive MIMO

ในระบบ Massive Multiple Input Multiple Output systems : Massive MIMO ปัจจุบันนั้นได้รับการพัฒนาและมีการพูดถึงกันอย่างแพร่หลาย เพราะเป็นเทคโนโลยีที่จะเข้ามามีบทบาทสำคัญต่อการใช้งานในอนาคต [1] งานวิจัยส่วนใหญ่เกี่ยวกับ Massive MIMO ยังตั้งอยู่บนสมมุติฐาน โดยที่การใช้เสาอากาศที่สถานีฐาน (base station) จำนวนหนึ่งซึ่งเกินจำนวนผู้ใช้ที่ใช้งานอยู่อย่างมาก การประมวลผลเชิงเส้น (linear processing) นั้นอาจจะเหมาะสมที่สุด (กับผู้ใช้เสาอากาศเดียว) นอกจากนี้ การใช้การรวมอัตราส่วนสูงสุด (MRC) หรือการส่งผ่านอัตราส่วนสูงสุด (MRT) ในอปลิงค์หรือดาวนลิงค์ ตามลำดับ ผลกระทบของสัญญาณรบกวนที่ไม่สัมพันธ์กันและการรบกวนภายในเซลล์มักจะหายไป เพราะตามที่กฎของตัวเลขจำนวนมากบอกเป็นนัยว่า เมทริกซ์ช่องสัญญาณ สำหรับผู้ใช้ที่ต้องการมีแนวโน้มที่จะตั้งฉากกับเมทริกซ์ช่องสัญญาณผู้ใช้ที่รบกวน การแสดงขั้นตอนมัลติเพล็กซ์เชิงพื้นที่อย่างง่ายพร้อมผลลัพธ์ที่ดีที่สุด Massive MIMO ยังให้ห้องปฏิบัติการที่มากขึ้น ซึ่งสามารถนำไปใช้เพื่อให้ RF ประหยัดพลังงานรวมไปถึงอัตราขยาย [2]-[3].

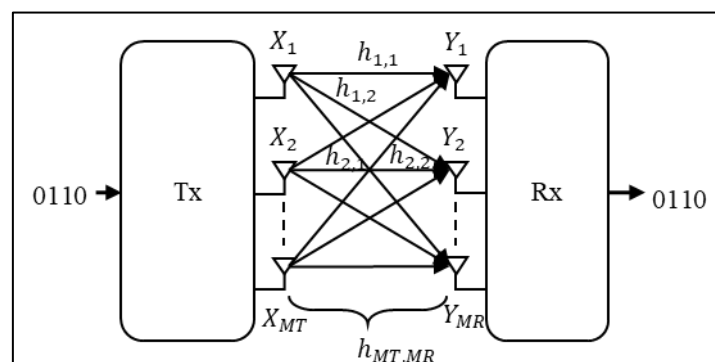
Massive MIMO ให้ประสิทธิภาพการใช้พลังงานในระดับสูง เมื่อเทียบกับระบบไร้สายแบบเดิม ดังที่แสดงในรูปที่ 2.1 ในดาวนลิงค์ สถานีฐานจะใช้การสร้างลำแสง ซึ่งได้รับการปรับปรุงด้วยจำนวนเสาอากาศที่เพิ่มขึ้น ส่งผลให้มีการส่งผ่านที่แม่นยำมากขึ้นในขณะที่ลดกำลังการแผ่รังสี นอกจากนี้ การเพิ่มจำนวนเสาอากาศที่สถานีฐานเป็นสองเท่า กำลังส่งสามารถลดลงได้ 3 เดซิเบล ทำให้ประสิทธิภาพโดยรวมเท่าเดิม ในสถานะที่เหมาะสมของการขยายพันธุ์และการประมวลผล ในการอปลิงค์ การสร้างบีมฟอร์มเมต (beamforming) ที่สอดคล้องกันจะทำให้ได้อาร์เรย์ที่สูงกว่าอีกครั้ง ทำให้กำลังส่งของผู้ใช้แต่ละรายลดลง ซึ่งสามารถใช้กับอุปกรณ์เคลื่อนที่ทั้งหมด [4]



รูปที่ 2.2 Spectral efficiency vs. Energy efficiency in SISO, MISO and Massive MIMO systems with different processing methods [4]

2.3 การประมาณช่องสัญญาณ (Channel Estimation)

ในการสื่อสารทั้งหมดสัญญาณจะผ่านสื่อที่เรียกว่าช่องสัญญาณ และถ้าหากสัญญาณเกิดการผิดเพี้ยนหรือมีเสียงรบกวนต่าง ๆ จะถูกเพิ่มเข้าไปในสัญญาณในขณะที่สัญญาณส่งผ่านเข้าสู่ช่องสัญญาณ ในการถอดรหัสสัญญาณที่ได้รับอย่างถูกต้องโดยไม่มีข้อผิดพลาดมากนักคือการลบความผิดเพี้ยนและเสียงรบกวนที่ใช้โดยช่องสัญญาณจากสัญญาณที่ได้รับ ซึ่งในการทำเช่นนี้มีขั้นตอนคือการหาลักษณะของช่องสัญญาณที่ส่งผ่านมาถึงแล้ว โดยมีเทคนิคหรือกระบวนการในการระบุลักษณะของช่องเรียกว่าการประมาณช่องสัญญาณ จะแสดงรายละเอียดดังนี้



รูปที่ 2.3 โครงสร้างระบบการสื่อสาร

จากรูปที่ 2.3 พิจารณาที่ X คือการกำหนดให้เป็นสัญญาณทางภาคส่ง (Tx) โดยที่ข้อมูลได้ผ่านกระบวนการเข้ารหัส (Modulation) จากนั้นจะถูกส่งไปที่ Y คือสัญญาณทางภาครับ (Rx) ซึ่งจะได้ช่องสัญญาณในรูปแบบ h_{M_R, M_T} โดยที่ M_R คือจำนวนสายอากาศภาครับและ M_T คือจำนวนสายอากาศภาคส่ง โดยที่สามารถเขียนอยู่ในรูปแบบสมการได้ดังนี้

$$Y = H \times X + n \quad (2.2)$$

โดยที่ n_{M_R} คือสัญญาณรบกวนที่สายอากาศภาคส่ง หากเขียนในรูปแบบแมทริกซ์จะอยู่ในรูปดังนี้

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{M_R} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1, M_T} \\ h_{12} & h_{22} & \cdots & h_{2, M_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R, 1} & h_{M_R, 2} & \cdots & h_{M_R, M_T} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{M_T} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{M_R} \end{bmatrix} \quad (2.3)$$

เมื่อพิจารณาที่การประมาณช่องสัญญาณ เราสามารถย้ายตัวแปรหาช่องสัญญาณที่ประมาณได้ต่อไปดังนี้

$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1, M_T} \\ h_{12} & h_{22} & \cdots & h_{2, M_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R, 1} & h_{M_R, 2} & \cdots & h_{M_R, M_T} \end{bmatrix} \quad (2.3)$$

$$\hat{H} = Y \times X^T \quad (2.4)$$

โดยที่ $(\circ)^T$ คือ conjugate transpose

ดังนั้นจากสมการที่ 2.4 เราสามารถหาการประมาณช่องสัญญาณเบื้องต้นได้ แต่เนื่องจากเทคนิคนี้เป็นเทคนิคพื้นฐาน ทางผู้วิจัยจึงอธิบายเพื่อให้เข้าใจถึงหลักการประมาณช่องสัญญาณเท่านั้นซึ่งเทคนิคในการประมาณช่องสัญญาณที่ใช้งานกันอย่างแพร่หลายคือ LS และ MMSE ดังนี้

2.3.1 เทคนิคการประมาณช่องสัญญาณ Least Square Estimation (LS)

เป็นแนวคิดช่วงแรก ๆ ในการที่จะประมาณช่องสัญญาณ โดยการประมาณช่องสัญญาณด้วยค่าตัวแปรต่าง ๆ ตามทฤษฎีจากสมการที่ 3.1 การประมาณค่าแบบนี้ถูกจัดประเภทเป็นตัวประมาณมิติเดียว (1D) ซึ่งหมายความว่า การประมาณช่องทำได้โดยใช้รอบการทดสอบ ในมิติเดียวไม่ว่าจะเป็นความถี่หรือเวลา การประมาณช่องสัญญาณ LS ช่วยลดปริมาณข้อผิดพลาดกำลังสองระหว่างสัญญาณที่ได้รับและค่าโดยประมาณ โดยสามารถอธิบายได้ว่า

$$\hat{H}_{LS} = \arg_{\tilde{H}_{LS}} \min \|Y - \tilde{H}_{LS} X\|^2 \quad (2.5)$$

โดยที่กำหนดให้ X คือเมทริกซ์ที่มีลำดับจำนวนตามสายอากาศภาคส่ง $X_{M_T} = 1, 2, 3, \dots, M_T$ ของจำนวนสายอากาศภาคส่งทั้งหมดดังนี้

$$X_N(p) = [X_1(p), \dots, X_{M_T}(p)]^T \quad (2.6)$$

ซึ่ง $(\circ)^T$ กำหนดให้เป็น conjugate transpose

โดยที่การประมาณช่องสัญญาณเป็นการตอบสนองระหว่างสายอากาศภาคส่งทั้งหมดและสายอากาศภาครับทั้งหมดตามจำนวน N -th กำหนดโดย

$$\hat{H}_{(N)LS} = YX^H (XX^H)^{-1} \quad (2.7)$$

โดยที่ $(\circ)^{-1}$ กำหนดเป็นเมทริกซ์ inverse หรือ inverse operation

หรือโดยทั่วไปสำหรับการประมาณช่องสัญญาณแบบไม่คำนึงถึง White Gaussian Noise กำหนดให้เป็น

$$\hat{H}_{(N)LS} = YR_{zz}^{-1} X^H (XR_{zz}^{-1} X^H)^{-1} \quad (2.8)$$

จาก R_{zz} คือค่า auto-correlation หรือสหสัมพันธ์อัตโนมัติแบบเมทริกซ์ของสัญญาณรบกวนโดยที่

$$R_{zz} = \sigma_n^2 I_{M_R \times M_R} \quad (2.9)$$

$I_{M_R \times M_R}$ คือ identity matrix หรือเมทริกซ์เอกลักษณ์

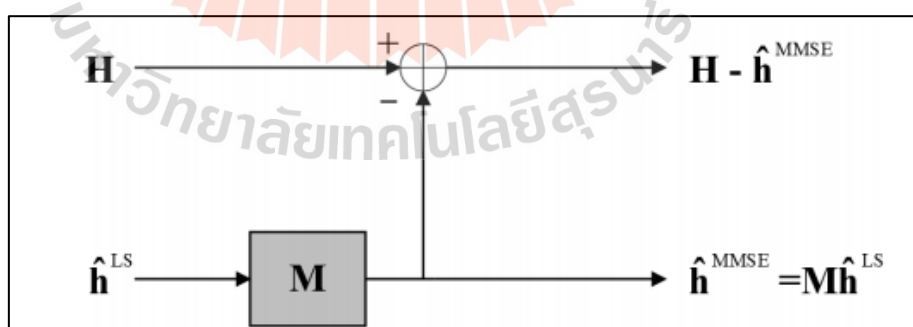
โดยที่ค่าความคลาดเคลื่อน (MSE) ของการประมาณช่องสัญญาณ LS เท่ากับ

$$MSE_{LS} = \frac{\sigma_n^2}{\sigma_x^2} \quad (2.10)$$

จากสมการที่ 2.10 ทำให้เห็นได้ชัดว่าค่าความคลาดเคลื่อนแปรผกผันกับ SNR โดยที่ $SNR = \sigma_x^2 / \sigma_n^2$ ซึ่งหมายความว่าตัวประมาณค่า LS มีความอ่อนไหวต่อการเพิ่มสัญญาณรบกวน โดยเฉพาะอย่างยิ่งเมื่อช่องสัญญาณอยู่ในระดับ deep fading อย่างไรก็ตามเนื่องจากเทคนิคนี้ไม่ได้พิจารณาพารามิเตอร์ทางสถิติจึงถูกนำมาใช้อย่างแพร่หลายสำหรับการประมาณช่องสัญญาณ [5]

2.3.2 เทคนิคการประมาณช่องสัญญาณ Minimum Mean-Square Error Channel Estimation (MMSE)

การประมาณช่องสัญญาณ MMSE เป็นเทคนิคที่แม่นยำมากขึ้นเมื่อเทียบกับการประมาณช่องสัญญาณแบบ LS โดยเราสามารถอธิบายการทำงานในรูปที่ โดยที่ M คือเมทริกซ์น้ำหนัก (weight matrix) และ \hat{H}_{MMSE} กำหนดให้เป็นช่องสัญญาณที่ประมาณได้ของ MMSE



รูปที่ 2.4 บล็อกไดอะแกรมของการประมาณช่องสัญญาณแบบ MMSE

การประมาณช่องสัญญาณแบบ MMSE นั้นจะลดขนาดของ MSE ระหว่างช่องสัญญาณจริง H และช่องสัญญาณ MMSE โดยประมาณช่องสัญญาณ \hat{H}_{MMSE} ด้วยวิธีการหา

ค่าประมาณเชิงเส้น (linear) ที่ดีที่สุดของการทำเมทริกซ์น้ำหนัก M และค่าประมาณช่องสัญญาณ LS มาคำนวณกัน โดยที่กำหนดให้

$$\hat{H}_{MMSE} = \arg_{\hat{H}_{MMSE}} \min \|H - \hat{H}_{MMSE}\|^2 \quad (2.11)$$

โดยที่

$$\hat{H}_{MMSE} = \hat{H}_{LS} M \quad (2.12)$$

พิจารณาที่ตามหลักการของมุมฉากเวกเตอร์ข้อผิดพลาดในการประมาณค่า $\varepsilon = H - \hat{H}_{MMSE}$ เป็นมุมฉากกับ \hat{H}_{LS} ส่งผลให้

$$0 = R_{H\hat{H}_{LS}} - MR_{\hat{H}_{LS}\hat{H}_{LS}} \quad (2.13)$$

โดยที่ $R_{H\hat{H}_{LS}}$ แทนค่าในเมทริกซ์สัมพันธ์ (Cross-correlation) ระหว่างช่องสัญญาณจริง H และ \hat{H}_{LS} ส่วน $R_{\hat{H}_{LS}\hat{H}_{LS}}$ คือการแทนค่าในเมทริกซ์สัมพันธ์แบบอัตโนมัติ (auto-correlation) ของค่า \hat{H}_{LS} กำหนดโดย

$$R_{\hat{H}_{LS}\hat{H}_{LS}} = R_{HH} + \frac{\sigma_n^2}{\sigma_x^2} I \quad (2.14)$$

ส่งผลกับสมการที่ 2.12 ในฟังก์ชัน M ผลลัพธ์ที่ได้คือ

$$M = R_{H\hat{H}_{LS}} R_{\hat{H}_{LS}\hat{H}_{LS}}^{-1} \quad (2.15)$$

จากนั้นทำการรวมสมการกันระหว่าง 2.15 และ 2.14 โดยสมการที่ 2.12 จะเขียนใหม่ได้เป็น

$$\hat{H}_{MMSE} = \hat{H}_{LS} R_{H\hat{H}_{LS}} \left(R_{HH} + \frac{\sigma_n^2}{\sigma_x^2} I \right)^{-1} \quad (2.16)$$

สมมติว่าพลังงานตอบสนองทุกช่องเป็นปกติเช่น $E\{|h_{M_R, M_T}|^2\} = \sigma_h^2$ วิธีแก้ปัญหามาโดยการประมาณของสมการ (2.16) สำหรับสายอากาศภาครับสัญญาณ N -th สามารถทำได้ง่ายขึ้นเป็น:

$$\hat{H}_{(N)MMSE} = YX^H \left(XX^H + \frac{\sigma_n^2}{\sigma_h^2} I \right)^{-1} \quad (2.17)$$

โดยสมการทั่วไปถ้าไม่พิจารณาที่ white Gaussian noise

$$\hat{H}_{(N)MMSE} = YR_{zz}^{-1} X^H \left(XR_{zz}^{-1} X^H + \frac{\sigma_n^2}{\sigma_h^2} I \right)^{-1} \quad (2.18)$$

เนื่องจากการประมาณช่องสัญญาณแบบ MMSE อาศัยการย่อขนาด MSE จึงมีประสิทธิภาพที่ดีกว่า การประมาณช่องสัญญาณแบบ LS โดยมีข้อเสียอยู่ที่ความจริงที่ว่ามันขึ้นอยู่กับสถิติของช่อง ดังนั้นวิธีนี้จึงมีความซับซ้อนสูงกว่าเทคนิคการประมาณช่องสัญญาณแบบ LS เพื่อลดความซับซ้อนของการประมาณช่องสัญญาณแบบ MMSE แนะนำให้ใช้เทคนิคที่เรียกว่า MMSE ที่แก้ไขแล้วใน [6] ดังนั้นทางผู้วิจัยจึงทำการเลือกการประมาณช่องสัญญาณที่กล่าวมาทั้ง 2 คือ LS และ MMSE นำมาเป็นตัวเปรียบเทียบกับเทคนิคที่ทางผู้วิจัยได้ทำการศึกษาอยู่

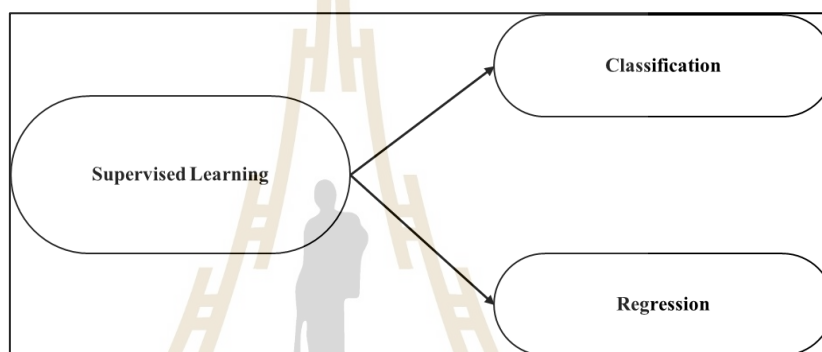
2.4 การเรียนรู้ของเครื่อง (Machine Learning ; ML)

ศาสตร์ด้านการเรียนรู้ของเครื่องเติบโตไปพร้อม ๆ กับปัญญาประดิษฐ์ ในความจริงนั้น การเรียนรู้ของเครื่องมีมาตั้งแต่ยุคแรก ๆ ของปัญญาประดิษฐ์ นักวิทยาศาสตร์หลายคนสนใจการสร้างเครื่องจักรที่สามารถเรียนรู้จากข้อมูลได้ จึงเริ่มทดลองวิธีการหลายๆอย่าง ที่เด่นชัดที่สุดคือโครงข่ายประสาทเทียม และในเวลาต่อมา ได้มีการคิดค้นโมเดลเชิงเส้นทั่วไปจากหลักการทางสถิติศาสตร์ ไปจนถึงการพัฒนาวิธีการให้เหตุผลตามหลักความน่าจะเป็น โดยเฉพาะในการประยุกต์ด้านการวินิจฉัยโรคอัตโนมัติ ซึ่งสามารถแบ่งประเภทของการเรียนรู้ของเครื่องโดยกว้าง ๆ

ได้หลายประเภทตามประเภทของข้อมูลนำเข้า (Input) หรือ ข้อมูลชุดฝึก และประเภทของของดังต่อไปนี้

2.4.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

Supervised Learning หรือการเรียนรู้แบบมีผู้สอนนั้น คือการทำให้คอมพิวเตอร์สามารถหาคำตอบของปัญหาได้ด้วยตัวเอง หลังจากเรียนรู้จากชุดข้อมูลตัวอย่างไปแล้วระยะหนึ่ง หลักการ Supervised Learning สามารถนำไปประยุกต์ได้ 2 รูปแบบคือ Classification และ Regression



รูปที่ 2.5 ประเภทของ Supervised Learning

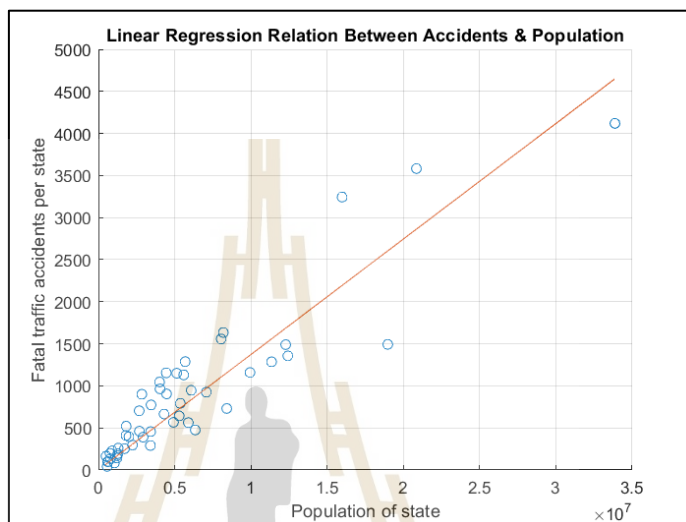
เป็นกลุ่มของ algorithm ที่ทำการสอนคอมพิวเตอร์โดยการศึกษาจากข้อมูลตัวอย่าง เพื่อให้คอมพิวเตอร์สามารถหาคำตอบของปัญหาได้ด้วยตัวเองหรือการแก้ปัญหา หลังจากเรียนรู้จากชุดข้อมูลตัวอย่างที่ได้ป้อนให้ไปแล้วระยะหนึ่ง

2.4.1.1 การวิเคราะห์การถดถอย (Regression)

โดยการวิเคราะห์การถดถอยอย่างง่าย (Simple Regression Analysis) เป็นการศึกษาระหว่างตัวแปร 2 ตัวในงานวิจัยนี้กำหนดให้เป็น X และ Y ที่มีความสัมพันธ์ในลักษณะเชิงเส้น (Linear) ดังนี้

$$Y_{mean} = a + bX \quad (2.19)$$

โดยที่ Y_{mean} เป็นค่าเฉลี่ยของ Y ทั้งหมด เนื่องจากการวิเคราะห์การถดถอยอย่างง่ายนั้น ตัวแปร X จะถูกกำหนดค่าไว้ก่อนและค่าของ Y จะเปลี่ยนแปลงตามตัวแปร X เนื่องจากค่า X ค่าหนึ่งจะมีค่า Y เป็นคู่ของค่า X ในหลายๆ ค่าและเมื่อนำค่า X และ Y มาทำการพล็อตกราฟแสดงความสัมพันธ์ระหว่างค่าเฉลี่ยซึ่งก็คือ เส้นกราฟถดถอย (Regression Line)



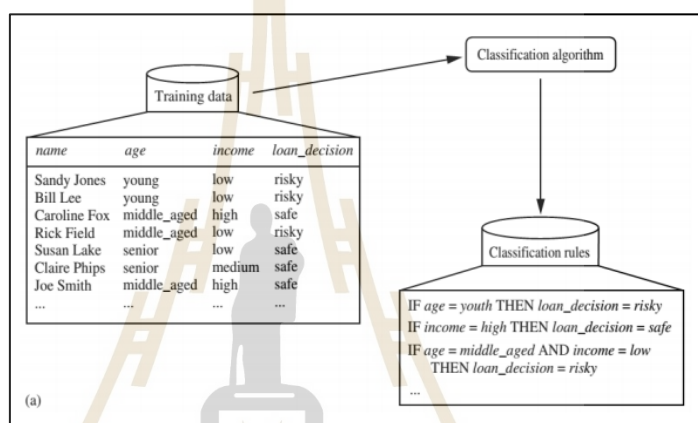
รูปที่ 2.6 ตัวอย่างกราฟระหว่างอุบัติเหตุและจำนวนประชากร

โดยจัดให้ Y คือจำนวนอุบัติเหตุ และ X คือจำนวนประชากรโดยที่ข้อมูลเป็นการยกตัวอย่างขึ้นมาเพื่อแสดงให้เห็นถึงกราฟความสัมพันธ์ถดถอย (Regression Line) ซึ่งจากสมการเส้นตรง $Y_{mean} = a + bX$ พิจารณาที่ a และ b เป็นตัวแปรที่ยังไม่ทราบค่า จึงต้องประมาณค่าโดยกลุ่มข้อมูลตัวอย่าง โดยวิธีที่นิยมใช้งานคือกำลังสองน้อยสุด Least Square Method (LS) ซึ่งนิยามค่า a คือค่าคงที่ (Constant) และค่า b คือค่าความชัน (Slop) ของเส้นกราฟ ส่งผลโดยตรงกับเส้นกราฟที่มีการเปลี่ยนแปลงระหว่าง Y และ X ทางผู้วิจัยเรียกส่วนนี้ว่า สัมประสิทธิ์การถดถอย (Regression Coefficient) หรือสามารถเรียกอีกแบบว่าสัมประสิทธิ์พยากรณ์

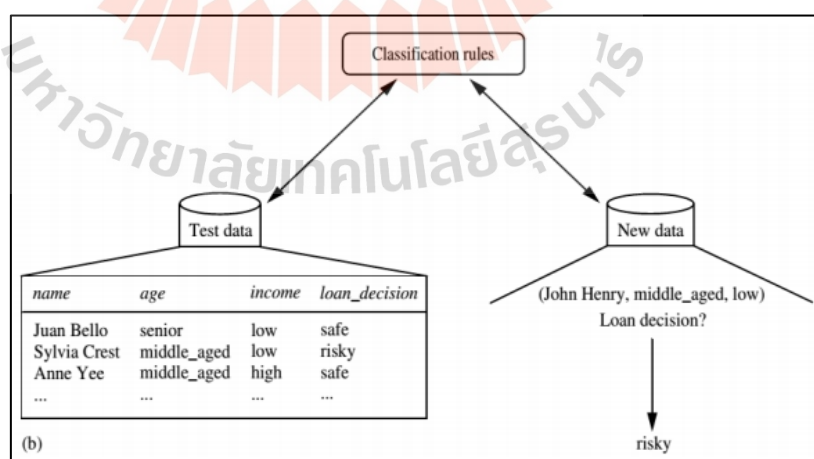
2.4.1.2 การจำแนก (classification)

คือการแบ่งข้อมูลออกเป็นหมวดหมู่ โดยกระบวนการของ classification แบบทั่วไปคือ การที่อัลกอริทึมพยายามจะประมาณการข้อมูล X ใน ชุดฝึกข้อมูล (Training set) และแสดงผล Y ออกมาในรูปแบบของหมวดหมู่ที่แบ่งแยกกันโดยชัดเจน (Discrete) ซึ่งสิ่งที่สำคัญอีกอย่างของการประมาณการคือ เราต้องมีการกำหนดเป้าหมาย (Target) ข้อมูลที่เราต้องการอะไร

โดยที่จำเป็นต้องเป็นแบบ ข้อมูลที่แยกกันอย่างชัดเจน เช่น เราต้องการทราบว่า นักเรียนห้อง A สอบผ่าน หรือสอบตก เมื่อเรานำข้อมูลของนักเรียนห้อง A มาผ่านการวิเคราะห์แบบ classification ผลที่ออกมา ก็จะเป็น 2 แบบ สอบผ่านและสอบตก หรือการอนุมัติสินเชื่อ เราก็อาจจะตั้งเป้าหมายว่า ผ่าน สินเชื่อหรือไม่ผ่านสินเชื่อ เป็นต้น แต่ทั้งนี้เป้าหมายของเราอาจจะไม่ได้แบ่งออกเป็นแค่ 2 แบบ ก็คือ ผ่านไม่ผ่าน แต่เราสามารถกำหนดเป้าหมายมากกว่า 2 อย่างก็ได้เช่นกัน เช่น แบ่งกลุ่มดอกไม้ ออกเป็น 4 กลุ่ม ตามตัวแปรที่เรียกว่าลักษณะเฉพาะ (features) ต่อไปนี้คือ สี ขนาดดอก แหล่งที่เกิด เป็นต้น



รูปที่ 2.7 การจำแนกข้อมูล (a)



รูปที่ 2.8 การเรียนรู้จากข้อมูลเพื่อสร้างตัวจำแนกข้อมูล (b)

ในตัวอย่างที่แสดงดังรูปที่ 2.7 และ 2.8 เป็นการกู้-ยืมเงิน โดยจากรูปที่ 2.7 จะเป็นกระบวนการสร้างตัวจำแนกข้อมูลจากชุดข้อมูลที่เป็นอินพุต ซึ่งแต่ละข้อมูลที่บันทึกจะประกอบไปด้วยเซตของแอทริบิวต์ที่บ่งบอกถึงคุณลักษณะของบุคคลที่กู้-ยืมเงิน และหมวดของบุคคลนั้น ๆ ว่ามีความปลอดภัยหรือมีความเสี่ยงในการกู้-ยืมเงินหรือไม่ โดยกระบวนการสร้างตัวจำแนกข้อมูลมักถูกเรียกว่า “learning” หรือ “training” ที่เกิดจากการนำเอาขั้นตอน วิธีสำหรับการจำแนกข้อมูลมาดำเนินการกับข้อมูล ข้อมูลที่บันทึก X หนึ่ง ๆ ในชุดข้อมูลที่ทำการพิจารณา ซึ่งจะประกอบไปด้วยเซตของแอทริบิวต์ $X = (x_1, x_2, \dots, x_n)$ ที่บ่งบอกถึงคุณลักษณะต่าง ๆ ของข้อมูลที่บันทึก นอกจากนั้นยังมีข้อมูลอีกหนึ่งแอทริบิวต์ที่บ่งบอกถึงหมวดหมู่ของข้อมูล (class label attribute) โดยแอทริบิวต์หมวดหมู่ของข้อมูลจะเป็นข้อมูลแบบไม่ต่อเนื่อง (discrete-valued) โดยชุดข้อมูลที่เป็นอินพุตสำหรับการสร้างตัวจำแนกชุดข้อมูลจะถูกเรียกว่า “ชุดข้อมูลสำหรับสอน (training data) ตัวอย่าง (sample/instances) ชุดข้อมูล (data points) หรือ สิ่งของ (objects)” เป็นต้น

ในขั้นตอนที่สองของการจำแนกข้อมูลดังรูปที่ 2.8 จะเป็นการเรียกใช้ตัวจำแนกข้อมูลที่ถูกสร้างขึ้นจากขั้นตอนที่หนึ่ง (รูปที่ 2.7) โดยในตอนเริ่มต้นตัวจำแนกข้อมูลจะถูกทดสอบและประเมินความถูกต้อง ถ้าเราใช้ชุดข้อมูลสำหรับสอนในการทดสอบตัวจำแนกข้อมูลจะทำให้ค่าความถูกต้องมีค่าค่อนข้างสูง เนื่องจากตัวจำแนกข้อมูลที่สร้างขึ้นจะเหมาะกับชุดข้อมูลนั้นเป็นอย่างมาก (overfit) แต่ถ้าเราใช้ชุดข้อมูลที่แตกต่างออกไปในการทดสอบ (test set) โดยชุดข้อมูลที่ใช้จะต้องมีแอทริบิวต์หมวดหมู่ข้อมูลแนบอยู่ด้วย จะทำให้เราทราบค่าความถูกต้องของตัวจำแนกข้อมูลได้ โดยค่าความถูกต้องของตัวจำแนกข้อมูลที่ถูกสร้างขึ้นจะเป็นเปอร์เซ็นต์ของตัวจำแนกข้อมูลที่สามารถจำแนกข้อมูลได้อย่างถูกต้อง (ตัวจำแนกข้อมูลบ่งบอกถึงหมวดหมู่ข้อมูลได้เหมือนกับหมวดหมู่ข้อมูลที่ถูกแนบมากับข้อมูลที่บันทึกหนึ่ง ๆ) เมื่อค่าความถูกต้องของตัวจำแนกข้อมูลมีค่าที่น่าพึงพอใจหรือยอมรับได้ เราจะใช้ตัวจำแนกข้อมูลในการจำแนกหรือบ่งบอกถึงหมวดหมู่ข้อมูลที่เข้ามาใหม่ที่เราไม่ทราบข้อมูลมาก่อน (ข้อมูลที่เข้ามาใหม่จะถูกเรียกว่า ‘unknown’ หรือ ‘previously unseen’ data) ตัวอย่างเช่น ตัวจำแนกข้อมูลถูกสร้างขึ้นในรูปที่ 2.7 จะถูกใช้เพื่อตัดสินใจให้กู้-ยืมเงินเอกสารที่ยื่นเข้ามาไม่ว่าจะให้กู้-ยืมหรือไม่ เป็นต้น

ซึ่งรูปที่ 2.7 และ 2.8 นั้นเป็นการทดสอบตัวจำแนกข้อมูลเพื่อวัดความถูกต้อง สำหรับการเปรียบเทียบประสิทธิภาพวิธีในการจำแนกข้อมูล มีเกณฑ์หลักที่ใช้อยู่ด้วยกันต่อไป

ตารางที่ 2.1 การเปรียบเทียบประสิทธิภาพวิธีในการจำแนกและทำข้อมูล

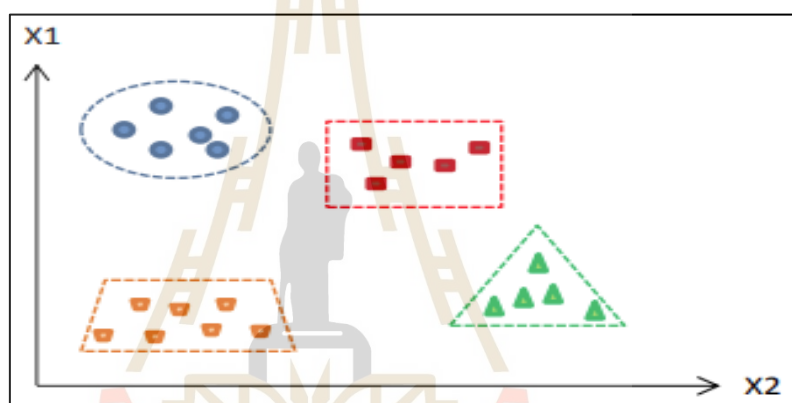
ความถูกต้อง (Accuracy)	จะเกี่ยวข้องกับความสามารถของตัวจำแนกข้อมูลที่ถูกสร้างขึ้นที่จะสามารถจำแนกข้อมูลที่ไม่เคยพบเจอมาก่อนได้อย่างถูกต้อง โดยในการวัดความถูกต้องอาจจะประเมินได้จากการใช้ชุดข้อมูลหนึ่ง ๆ หรือมากกว่า ที่แยกจากชุดข้อมูลเรียนรู้ (training dataset)
ความเร็ว (Speed)	จะเกี่ยวข้องกับเวลาที่ใช้ในการคำนวณทั้งในส่วนของการสร้างตัวจำแนกข้อมูลและการจำแนกหรือทำนายข้อมูล
ความทนทาน (Robustness)	จะเกี่ยวข้องกับความสามารถของตัวจำแนกหรือตัวทำนายข้อมูลที่จะทำการทำนายได้อย่างถูกต้องจากข้อมูลตั้งต้นที่มีสิ่งรบกวนหรืออิทธิกรรขาดหายไปของข้อมูล
ความยืดหยุ่นต่อปริมาณข้อมูล (Scalability)	จะเกี่ยวข้องกับความสามารถในการสร้างตัวจำแนกข้อมูลหรือตัวทำนายข้อมูลได้อย่างมีประสิทธิภาพเมื่อมีข้อมูลที่ต้องพิจารณาเป็นปริมาณมาก
ความสามารถในการเข้าใจ (Interpretability)	เกี่ยวข้องกับระดับความสามารถที่จะถูกเข้าใจในตัวจำแนกหรือทำนายข้อมูลจากผู้ใช้งาน

2.4.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

ในการเรียนรู้แบบมีผู้สอน เรารู้ได้ว่าคำตอบที่ควรจะเป็นคืออะไร และเราก็ระบุคำตอบที่ถูกต้องลงไป ใน Label แต่ในการเรียนรู้แบบไม่มีผู้สอน เราจะไม่มีคำตอบที่แน่ชัด แต่เราต้องการให้เครื่องตามหาโครงสร้างข้อมูลที่เราไม่รู้จัก (Unknown Structure) เช่น ในโครงสร้าง DNA เราอยากตามหาชิ้นที่ส่งผลให้คนมีตาสีฟ้าหรือสีดำ เราไม่รู้ว่าจะตรงไหนของ DNA ที่ส่งผลต่อสีตาเรารู้แต่ว่าข้อมูล DNA ของคนที่เรานำมาทดสอบมีสองกลุ่ม กลุ่มหนึ่งคือ คนที่มีตาสีฟ้า อีกกลุ่มคือคนที่มีตาสีดำ ดังนั้นเราอยากให้เครื่องตรวจหาว่าส่วนไหนของ DNA ที่แสดงความแตกต่างออกมาอย่างเด่นชัดระหว่างกลุ่ม และเหมือนกันมากในกลุ่มเดียวกัน โดยที่การเรียนรู้แบบไม่มีผู้สอนสามารถแบ่งออกได้เป็น 2 ประเภทคือ การจับกลุ่มของข้อมูล (Clustering), การลดจำนวนมิติเพื่อบีบอัดข้อมูล (Dimensionality Reduction) สามารถอธิบายได้ดังนี้

2.4.2.1 การจับกลุ่มของข้อมูล (Clustering)

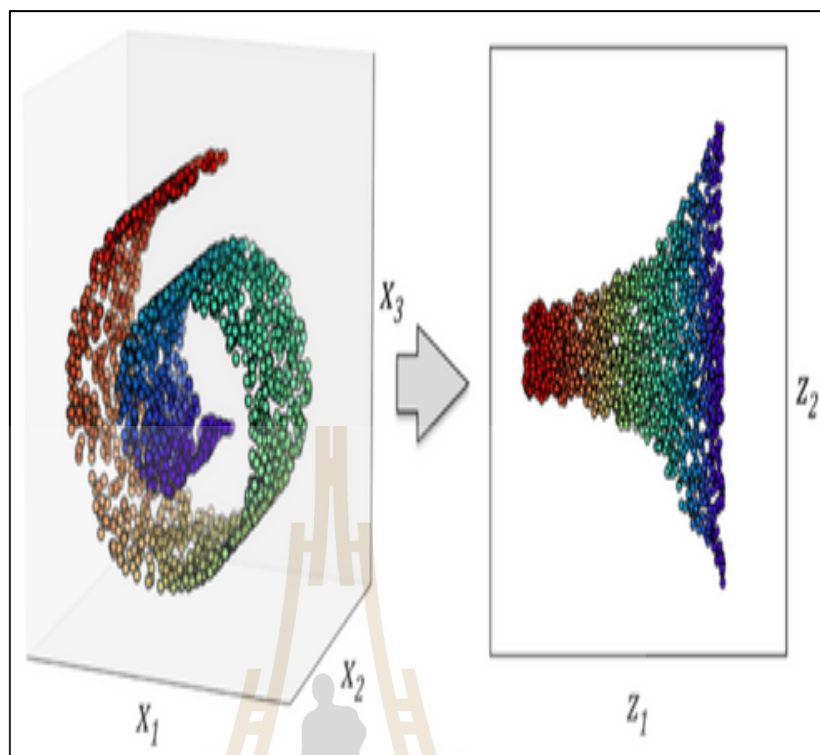
อัลกอริทึมจะตรวจสอบเฉพาะข้อมูลที่ป้อนเข้ามาเท่านั้น โดยปราศจากการให้ผลลัพธ์ที่จะเกิดขึ้น เช่น การสำรวจข้อมูลประชากรเพื่อหาแบบแผน (Pattern) ของข้อมูลนั้น เราต้องทราบว่าข้อมูลของเราประกอบไปด้วย Class เท่าใด โดยเราต้องระบุว่าแต่ละกลุ่มที่เครื่องมือพบคืออะไร จริงทำให้เราสามารถลดจำนวนของ Label ที่เราต้องระบุให้เหลือจำนวนเท่ากับ class ได้ ตัวอย่างเช่น มีการจับกลุ่มข้อมูลเข้าสองมิติ (ข้อมูลเข้าเป็นตัวเลข 2 ค่า คือ x_1 และ x_2) โดยยังไม่ทราบว่าข้อมูลดังกล่าวอยู่กลุ่มใดแต่เรากำหนดให้เครื่องมือหาวิธีออกจากกันได้จากตัวอย่างแบบกลุ่มออกเป็นสี่กลุ่มและได้ผลออกมาเป็นขอบเขตของแต่ละกลุ่ม ดังรูปที่ภาพต่อไปนี้



รูปที่ 2.9 การแบ่งกลุ่มของการจับกลุ่มข้อมูล

2.4.2.2 อัดข้อมูล (Dimensionality Reduction)

เป็นการลดจำนวนมิติเพื่อบีบอัดข้อมูล ทำให้เราไม่จำเป็นต้องเก็บข้อมูลไม่ครบแต่ก็ยังสามารถจำแนกข้อมูลได้โดย โดยสิ่งที่เราต้องการคือการลดข้อมูลให้เหลือ 2 มิติและยังสามารถนำไปแยกประเภท class ได้คืออย่างเดิม



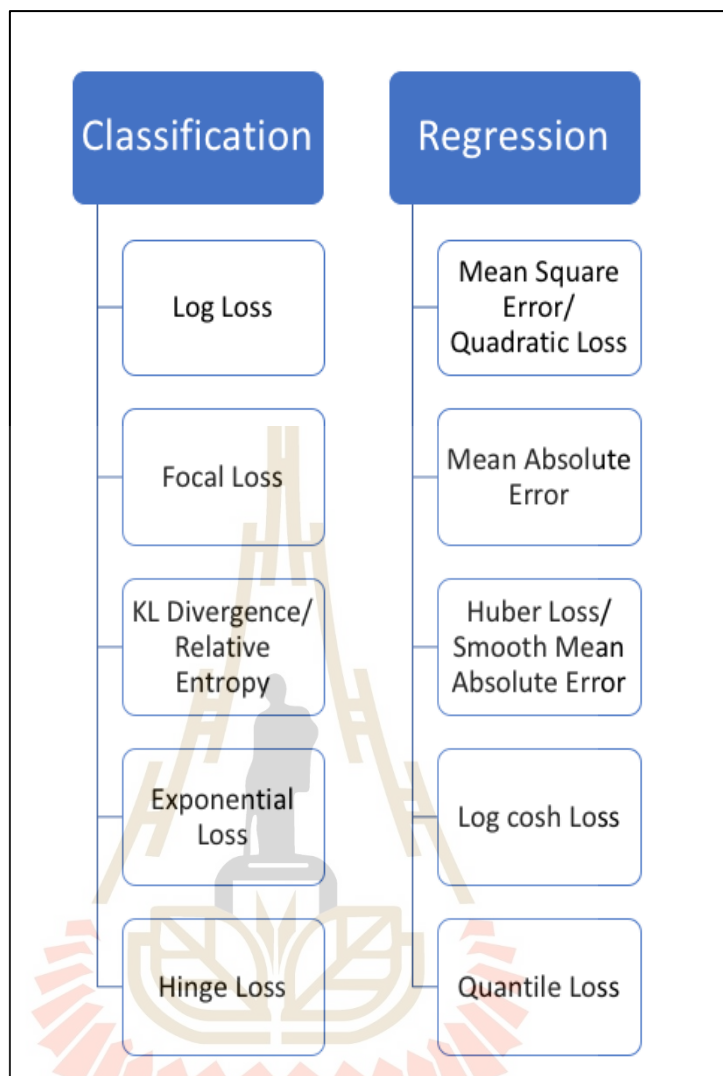
รูปที่ 2.10 การแสดงผลของการอัดข้อมูลแล้วนำไปแยกประเภท class [Figure credit: Sebastian Raschka, Python Machine Learning]

2.5 สิ่งที่จะวิเคราะห์เปรียบเทียบ

ในการวัดค่าประสิทธิภาพข้อมูลจากการประยุกต์ใช้การเรียนรู้ของเครื่องนั้นมีด้วยกันหลายวิธี เช่น การวัดเฉพาะระบบสื่อสารและการวัดค่าความเสถียรของข้อมูลที่ผ่านกระบวนการเรียนรู้ของเครื่องซึ่งมีดังต่อไปนี้

2.5.1 ค่าความคลาดเคลื่อนของข้อมูล (Loss function)

สำหรับการวัดค่าประสิทธิภาพข้อมูลจากการประยุกต์ใช้การเรียนรู้ของเครื่องมีด้วยกันหลากหลายวิธีดังรูป [7]



รูปที่ 2.11 กลุ่มเครื่องมือวัด Classification and Regression Loss.
(www.heartbeat.fritz.ai)

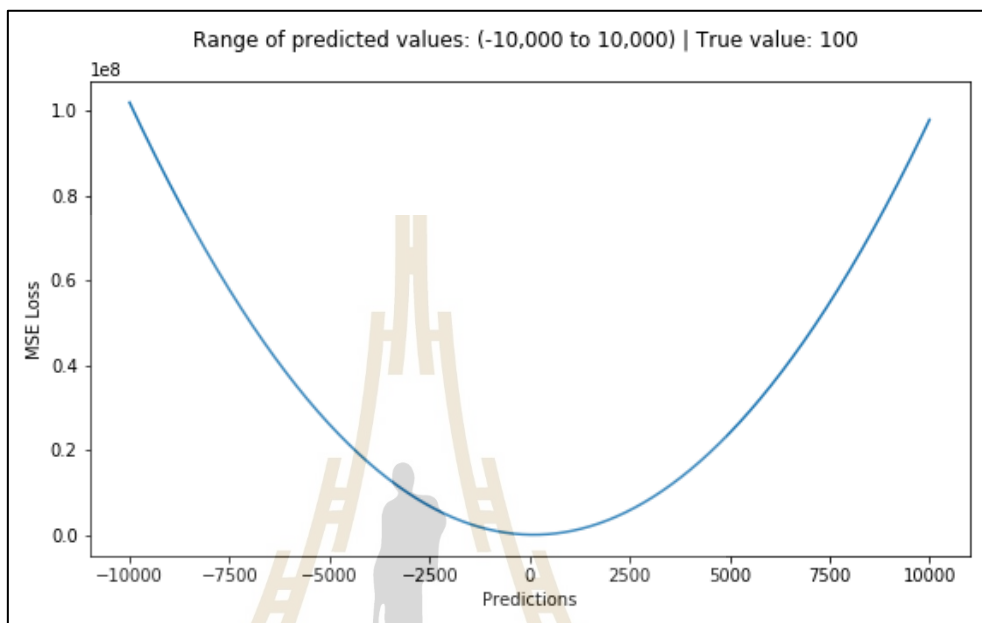
สำหรับการวัดค่าประเภท Regression loss นั้นมีดังนี้

- 1) Mean Square Error, Quadratic loss (L2 Loss)

เป็นฟังก์ชันการสูญเสียการถดถอยที่ใช้บ่อยที่สุด MSE คือผลรวมของค่าทั้งหมดยกกำลังสองระหว่างตัวแปรเป้าหมายกับค่าที่คาดการณ์ไว้

$$MSE = \frac{\sum_{i=1}^n (y_i - y_i^p)^2}{n} \quad (2.20)$$

ด้านล่างนี้คือพล็อตของฟังก์ชัน MSE โดยที่ค่าเป้าหมายที่แท้จริงคือ 100 และค่าที่คาดการณ์จะอยู่ในช่วงระหว่าง -10,000 ถึง 10,000 การสูญเสีย MSE (แกน Y) ถึงค่าต่ำสุดที่การคาดการณ์ (แกน X) = 100 ช่วงคือ 0 ถึง ∞

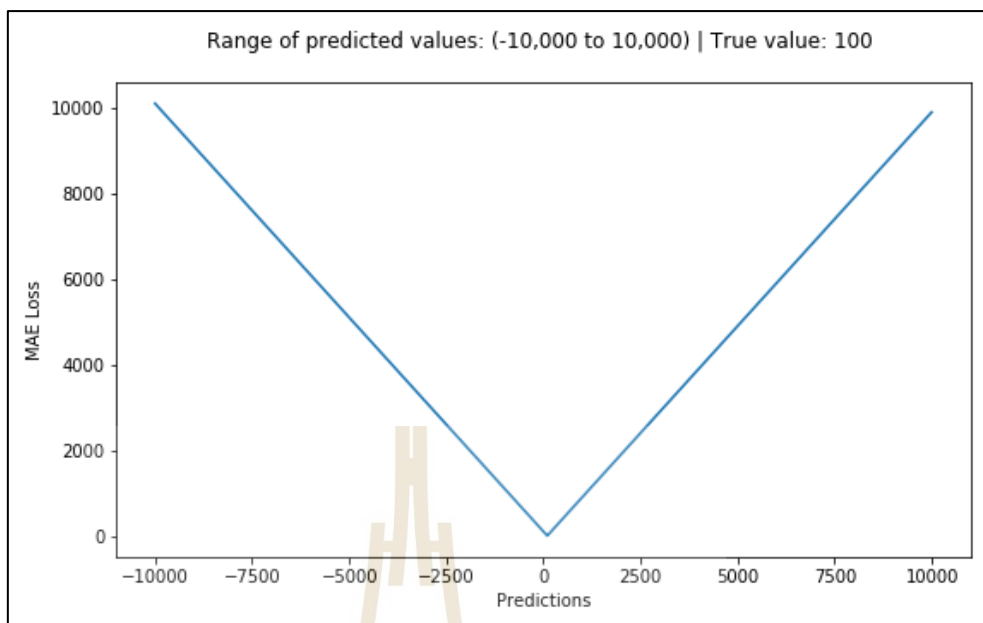


รูปที่ 2.12 กราฟของ MSE Loss (Y-axis) เทียบกับ Predictions (X-axis)

2) Mean Absolute Error, (L1 Loss)

เป็นฟังก์ชันการสูญเสียอีกรูปแบบที่ใช้สำหรับตัวแบบการถดถอย MAE คือผลรวมของความแตกต่างสัมบูรณ์ระหว่างเป้าหมายของเรากับตัวแปรที่คาดการณ์ไว้ ดังนั้นจึงวัดขนาดเฉลี่ยของข้อผิดพลาดในชุดการคาดคะเน โดยไม่คำนึงถึงทิศทาง หากเราพิจารณาทิศทางด้วย จะเรียกว่า Mean Bias Error (MBE) ซึ่งเป็นผลรวมของเศษต่อข้อผิดพลาด ช่วงยังเป็น 0 ถึง ∞

$$MSE = \frac{\sum_{i=1}^n |y_i - y_i^p|}{n} \quad (2.21)$$

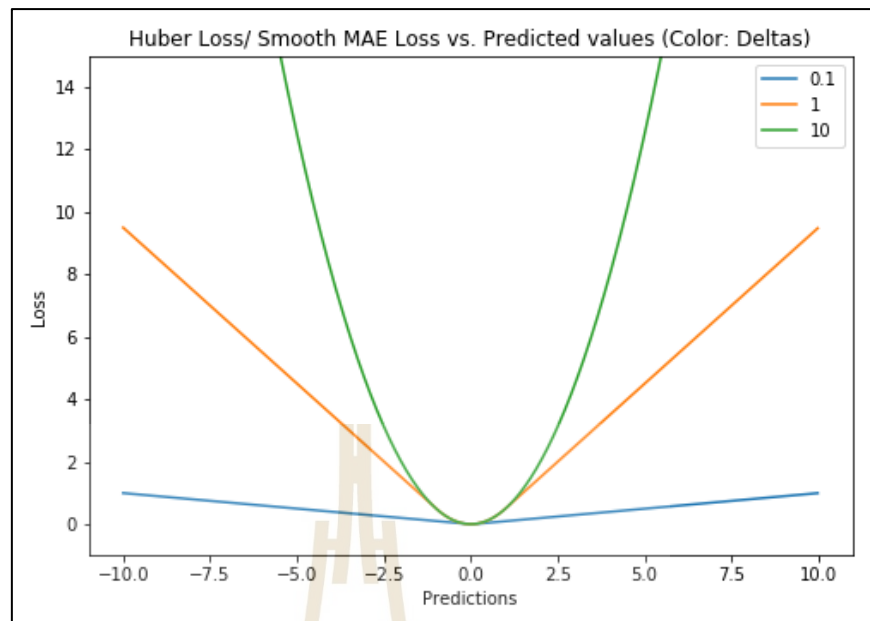


รูปที่ 2.13 กราฟของ MAE Loss (Y-axis) เทียบกับ Predictions (X-axis)

3) Huber Loss (Smooth Mean Absolute Error)

มีความอ่อนไหวต่อค่าผิดปกติในข้อมูลน้อยกว่าการสูญเสียข้อผิดพลาดกำลังสอง นอกจากนี้ยังสามารถหาอนุพันธ์ได้ที่ 0 โดยพื้นฐานแล้วมันเป็นข้อผิดพลาดแบบสัมบูรณ์ซึ่งจะกลายเป็นกำลังสองเมื่อข้อผิดพลาดมีขนาดเล็ก ความคลาดเคลื่อนนั้นมีขนาดเล็กเพียงใดในการทำให้มันเป็นกำลังสองนั้นขึ้นอยู่กับไฮเปอร์พารามิเตอร์ δ (เคลต้า) ซึ่งสามารถปรับได้ การสูญเสีย Huber เข้าใกล้ MSE เมื่อ $\delta \sim 0$ และ MAE เมื่อ $\delta \sim \infty$ ที่มีจำนวนมาก

$$L_{\delta}(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & \text{for } |y - f(x)| \leq \delta, \\ \delta|y - f(x)| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases} \quad (2.22)$$



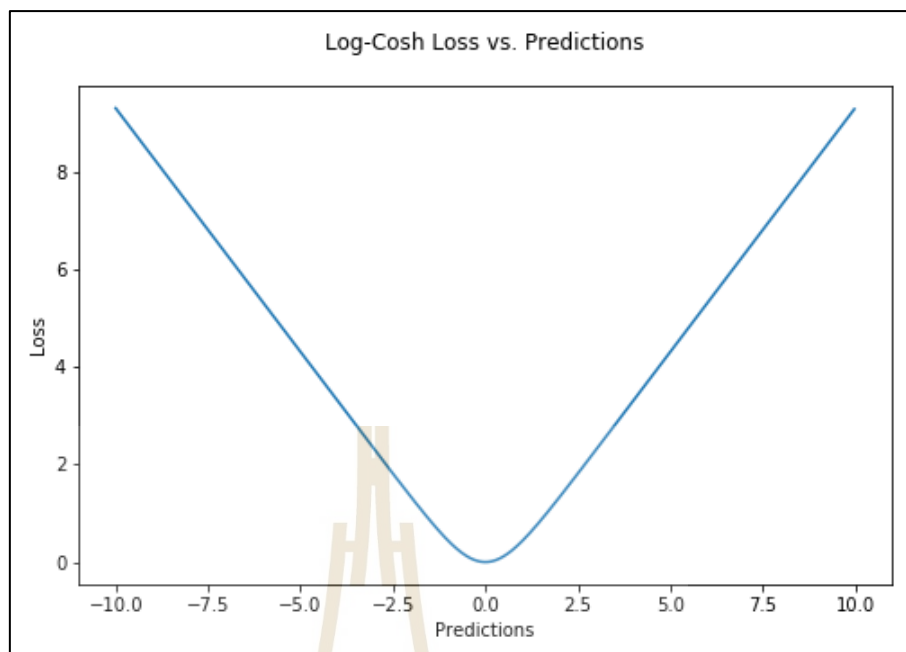
รูปที่ 2.14 กราฟของ Huber Loss (Y-axis) เทียบกับ Predictions (X-axis)

4) Log-Cosh Loss

Log-cosh เป็นฟังก์ชันอีกรูปแบบที่ใช้ในงานการถดถอยที่ราบรื่นกว่า L2 Log-cosh คือลอการิทึมของไฮเพอร์โบลิกโคไซน์ของข้อผิดพลาดในการทำนาย

$$L_y(y, y^p) = \sum_{i=1}^n \log(\cosh(y_i^p - y_i)) \quad (2.23)$$

ข้อดีคือ $\log(\cosh(x))$ มีค่าประมาณเท่ากับ $(x^2)/2$ สำหรับ x ขนาดเล็กและ $\text{abs}(x) - \log(2)$ สำหรับ x ขนาดใหญ่ ซึ่งหมายความว่า 'logcosh' ทำงานเหมือนกับค่าคลาดเคลื่อนกำลังสองเฉลี่ยเป็นส่วนใหญ่ แต่จะไม่ได้รับผลกระทบอย่างมากจากการคาดการณ์ที่ไม่ถูกต้องเป็นครั้งคราว มีข้อดีทั้งหมดของการสูญเสียของ Huber และสามารถแยกความแตกต่างได้สองเท่าทุกที่ซึ่งแตกต่างจากการสูญเสียของ Huber



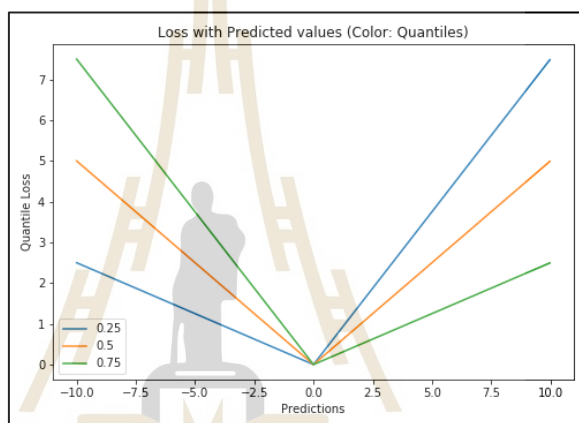
รูปที่ 2.15 กราฟ Log-cosh Loss (Y-axis) เทียบกับ Predictions (X-axis).

5) Quantile Loss

ปัญหาการทำนายผลในความเป็นจริงส่วนใหญ่ จะให้ความสนใจเกี่ยวกับความไม่แน่นอนในการทำนายของเรา การรู้ช่วงของการคาดการณ์ซึ่งต่างจากการประมาณการแบบจุดเท่านั้นสามารถปรับปรุงกระบวนการตัดสินใจสำหรับปัญหาทางธุรกิจจำนวนมากได้อย่างมีนัยสำคัญ Quantile Loss กลายเป็นประโยชน์เมื่อเราสนใจในการทำนายช่วงเวลาแทนที่จะเป็นเพียงการคาดคะเนแบบจุด ช่วงการทำนายจากการถดถอยกำลังสองน้อยที่สุดขึ้นอยู่กับสมมติฐานว่าเศษเหลือ (ค่าจริง — ค่าประมาณ) มีความแปรปรวนคงที่ในค่าของตัวแปรอิสระ เราไม่สามารถเชื่อถือตัวแบบการถดถอยเชิงเส้นที่ละเมิดสมมติฐานนี้ได้ เราไม่สามารถทิ้งแนวคิดเรื่องการปรับโมเดลการถดถอยเชิงเส้นให้เป็นพื้นฐาน โดยบอกว่าสถานการณ์ดังกล่าวจะเป็นแบบจำลองที่ดีกว่าเสมอโดยใช้ฟังก์ชันที่ไม่เป็นเชิงเส้นหรือแบบจำลองแบบต้นไม้ นี่ก็จุดที่การสูญเสียเชิงปริมาณและการถดถอยเชิงปริมาณเข้ามาช่วย เนื่องจากการถดถอยตามการสูญเสียเชิงปริมาณให้ช่วงเวลาการคาดการณ์ที่สมเหตุสมผลแม้สำหรับส่วนที่เหลือที่มีความแปรปรวนไม่คงที่หรือการแจกแจงแบบไม่ปกติ การถดถอยตามควอนไทล์มีจุดมุ่งหมายเพื่อประเมิน "ปริมาณ" แบบมีเงื่อนไขของตัวแปรตอบสนอง โดยให้ค่าบางอย่างของตัวแปรทำนาย การสูญเสียเชิงปริมาณเป็นเพียงส่วนเสริมของ MAE (เมื่อควอนไทล์เป็นเปอร์เซ็นต์ไทล์ที่ 50 มันคือ MAE) แนวคิดคือการเลือกค่าควอนไทล์โดย

พิจารณาจากว่าเราต้องการให้ค่ากับข้อผิดพลาดเชิงบวกหรือข้อผิดพลาดเชิงลบมากขึ้น ฟังก์ชันการสูญเสียพยายามให้บทลงโทษที่แตกต่างกันสำหรับการประเมินค่าสูงไปและการประเมินค่าต่ำไป โดยพิจารณาจากมูลค่าของควอนไทล์ที่เลือก (γ) ตัวอย่างเช่น ฟังก์ชันการสูญเสียเชิงปริมาณของ $\gamma = 0.25$ ให้โทษมากกว่าสำหรับการประเมินค่าสูงไป และพยายามรักษาค่าการทำนายให้ต่ำกว่าค่ามัธยฐานเล็กน้อย

$$L_\gamma(y, y^p) = \sum_{i=y_i < y_i^p} (\gamma - 1) \cdot |y_i - y_i^p| + \sum_{i=y_i > y_i^p} (\gamma) \cdot |y_i - y_i^p| \quad (2.24)$$

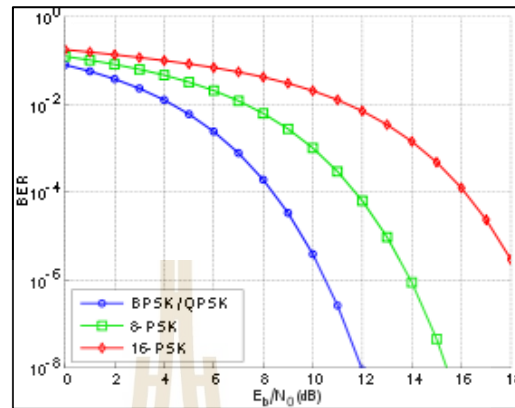


รูปที่ 2.16 กราฟ Log-cosh Loss (Y-axis) เทียบกับ Predictions (X-axis).

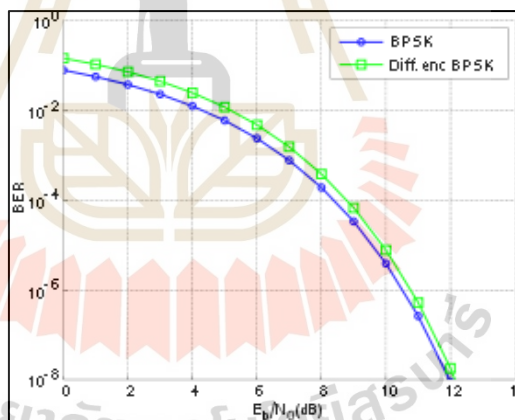
2.5.2 อัตราบิดผิดพลาด (BER)

ในการส่งผ่านดิจิทัลจำนวนของข้อผิดพลาดบิดคือจำนวนที่ได้รับบิดของกระแสข้อมูลกว่าช่องทางการสื่อสารที่มีการเปลี่ยนแปลงเนื่องจากเสียง , รบกวน , การบิดเบือนหรือบิด การประสานข้อผิดพลาด ประสิทธิภาพของระบบสื่อสาร แสดงด้วย อัตราบิดผิดพลาดหรืออัตราผิดพลาดในการส่งข้อมูล (Bit Error rate ; BER) ตัวอย่างเช่นสมมติว่าลำดับบิตที่ส่งนี้ 0 1 1 0 0 0 1 0 1 1 และลำดับบิตที่ได้รับต่อไปนี้ 0 0 1 0 1 0 1 0 0 1 จำนวนข้อผิดพลาดบิต (บิตที่ขีดเส้นใต้) คือในกรณีนี้ 3 BER คือ 3 บิตที่ไม่ถูกต้องหารด้วย 10 บิตที่ถ่ายโอนทำให้ได้ BER 0.3 หรือ 30% การวิเคราะห์ BER อาจได้รับการประเมินโดยใช้การจำลองทางคอมพิวเตอร์แบบสุ่ม หากสมมติว่ามีการใช้โมเดลช่องทางการส่งและแบบจำลองแหล่งข้อมูลอย่างง่าย BER อาจถูกคำนวณในเชิงวิเคราะห์ด้วย ตัวอย่างของแบบจำลองแหล่งข้อมูลดังกล่าวคือแหล่งข้อมูลBernoulli ตัวอย่างของแบบจำลองช่องสัญญาณอย่างง่ายที่ใช้ในทฤษฎีได้แก่ ช่องสัญญาณสมมาตรแบบไบนารี และ ช่องสัญญาณ

รบกวน Gaussian (AWGN) สถานการณ์ที่เลวร้ายที่สุดคือช่องสัญญาณสุ่ม โดยสิ้นเชิงซึ่งสัญญาณรบกวนจะครอบงำสัญญาณที่เป็นประโยชน์โดยสิ้นเชิง สิ่งนี้ส่งผลให้การส่ง BER เป็น 50%



รูปที่ 2.17 อัตราบิดผิดพลาดสำหรับ BPSK , QPSK , 8-PSK และ 16-PSK, โดยใช้ช่องสัญญาณ AWGN



รูปที่ 2.18 อัตราบิดผิดพลาดสำหรับ BPSK และ BPSK ที่แตกต่างกันโดยใช้ช่องสัญญาณ AWGN

ในช่องสัญญาณที่มีเสียงดัง BER มักจะแสดงเป็นฟังก์ชันของการวัดอัตราส่วนของพาหะต่อสัญญาณรบกวนที่เป็นมาตรฐานซึ่งแสดงถึง E_b/N_0 (อัตราส่วนพลังงานต่อบิตต่อความหนาแน่นสเปกตรัมของพลังงานเสียง) หรือ E_s/N_0 (พลังงานต่อสัญลักษณ์การมอดูเลตเป็นความหนาแน่นของสัญญาณรบกวน) ตัวอย่างเช่นในกรณีของการมอดูเลต QPSK และช่อง AWGN ฟังก์ชัน BER ตามของ E_b/N_0 จะถูกกำหนดโดย

$$BER = \frac{1}{2} \operatorname{erfc}\left(\sqrt{E_b / N_0}\right) \quad (2.25)$$

การพล็อตเส้นโค้ง BER เพื่ออธิบายประสิทธิภาพของระบบสื่อสารดิจิทัล ในการสื่อสารด้วยแสงมักใช้ BER (dB) เทียบกับกำลังรับ (dBm) ในขณะที่อยู่ในการสื่อสารแบบไร้สายจะใช้ BER (dB) กับ SNR (dB) [8]

2.5.3 ความจุช่องสัญญาณในระบบ MIMO (MIMO Channel Capacity)

หัวข้อนี้เสนอความจุช่องสัญญาณโดยทฤษฎีของ Shannon ซึ่งจะให้อัตราการส่งข้อมูล สูงสุด ภายใต้ช่องสัญญาณที่มีความน่าจะเป็นในการเกิดความผิดพลาดน้อยความจุช่องสัญญาณ เทียบกับปริมาณที่สูญเสียอธิบายโดยอัตราเร็วการส่งข้อมูล ได้จากการส่งผ่านช่องสัญญาณ ซึ่งมีความน่าจะเป็นในการเกิดความผิดพลาดไม่เป็นศูนย์ความจุช่องสัญญาณอยู่ภายใต้การรับรู้ สถานะช่องสัญญาณ รวมถึงอัตราขยายช่องสัญญาณทั้งภาคส่งและภาครับ ในส่วนแรกจะอธิบายถึง ความจุช่องสัญญาณที่ไม่มีการเปลี่ยนแปลงภายใต้ความแตกต่างในการสมมติช่องสัญญาณที่รับรู้ได้

เมื่อไม่มีการรับรู้สถานะข้อมูลที่ภาคส่งความจุช่องสัญญาณในระบบ MIMO ที่ใช้การประมวลผลเมมเบร่าลำดับ แสดงได้ดังนี้

$$C = \log_2 \det \left[I_{M_R} + \frac{P_t}{P_n M_T} H H^H \right] \quad (2.26)$$

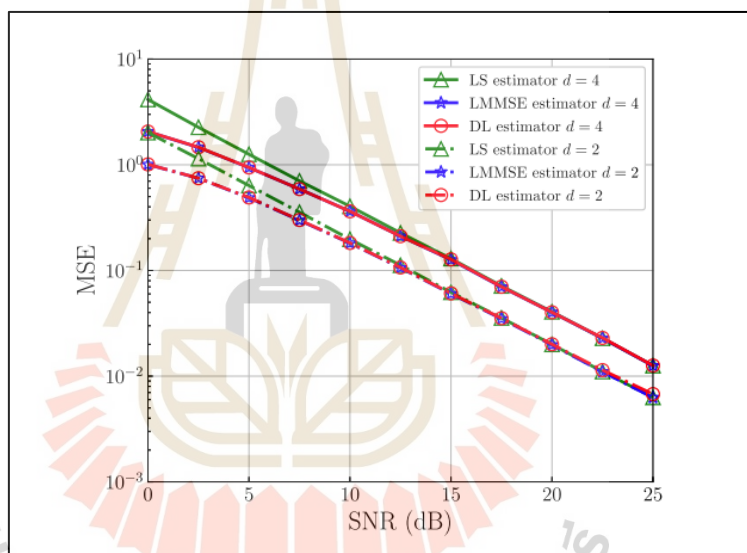
โดยที่ (2.26) มีหน่วยเป็นบิตต่อวินาทีต่อเฮิรตซ์เมื่อ I_{M_R} คือเมทริกซ์เอกลักษณ์ขนาด $M_R \times M_T$ ส่วน H คือช่องสัญญาณ ขนาด $M_R \times M_T$ และ H^H คือการทรานสโพสคอนจูเกตของเมทริกซ์ช่องสัญญาณ และ P_t / P_n คืออัตราส่วนสัญญาณที่รับได้ต่อสัญญาณรบกวน [9]

2.6 ปรัชญ่วรรณกรรมและงานวิจัย

ในส่วนนี้จะเป็นการกล่าวถึงงานวิจัยที่ทางผู้วิจัยได้ทำการศึกษาและค้นคว้าหาข้อมูลเพื่อนำมาประกอบงานวิจัยให้มีความต่อเนื่องและนำมาปรับปรุงพัฒนาผลลัพธ์ของงานวิจัยชิ้นนี้ โดยในงานวิจัยต่อไป [17] เป็นการกล่าวถึงอัลกอริทึมในกลุ่มการเรียนรู้เชิงลึกหรือเรียกว่า

(Deep Learning; DL) โดยเป้าหมายคือการเพิ่มประสิทธิภาพในการหาผลลัพธ์ของการประมาณช่องสัญญาณซึ่งในงานวิจัยชิ้นนี้เป็นการพิจารณาในระบบ Single Input Multiple Output ในระบบการสื่อสารสำหรับการประมาณช่องสัญญาณ โดยสมมุติให้มีการกำหนดสายอากาศที่สถานีฐาน (Base Station) และจำนวน ผู้ใช้งาน โดยรอบ ซึ่งได้กำหนดค่าต่าง ๆ ตามความเป็นจริง โดยในงานวิจัยชิ้นนี้ใช้วิธีการคำนวณหาค่าประมาณช่องสัญญาณ โดยเทคนิค LS, MMSE เทียบกับ DL โดยทำการประเมินประสิทธิภาพแบบหาค่าความคลาดเคลื่อนผิดพลาด MSE ซึ่งเป็นที่นิยมในการหาความถูกต้องของงาน

โดยแสดงผลลัพธ์ของงานวิจัยชิ้นนี้ดังรูปภาพที่ 2.19

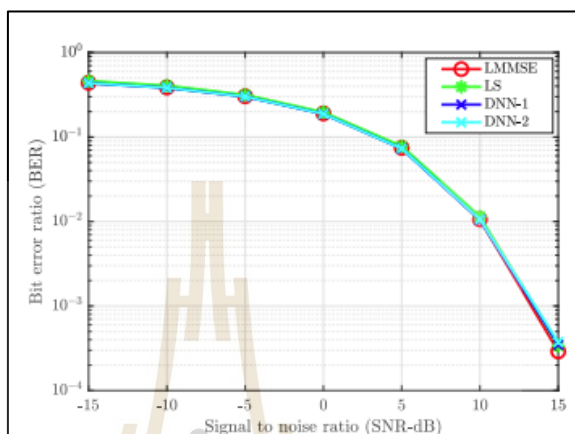


รูปที่ 2.19 ประสิทธิภาพ MSE โดยเทคนิค LS, MMSE และ DL เทียบกับ SNR [17]

โดยในรูปที่ 2.19 d คือจำนวนของสายอากาศ ซึ่งในรูปได้พิจารณาที่ 2 และ 4 ตัวในการประเมินประสิทธิภาพ ทำให้เห็นว่าวิธีการของ DL นั้นมีประสิทธิภาพการทำงานที่ดีกว่า LS และ MMSE ในช่วง 0 SNR (dB) ถึง 10 SNR (dB) และคู่แข่งกันในช่วง 10 SNR (dB) ถึง 25 SNR (dB) โดยถือว่าเป็นผลลัพธ์ที่ยอมรับได้ในงานวิจัยชิ้นนี้

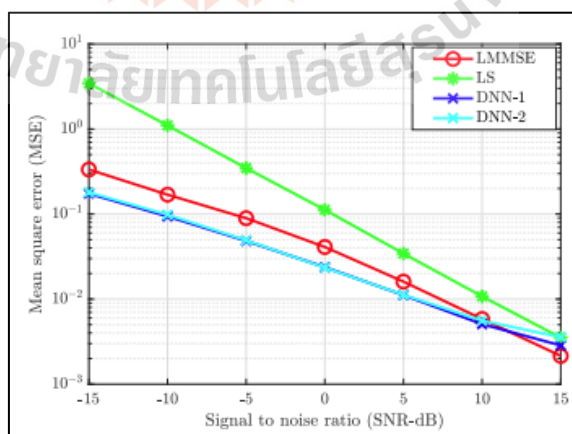
ต่อมาเป็นงานวิจัยที่ [18] เป็นงานวิจัยในระบบการสื่อสาร 5G โดยนำเสนอ DNN ที่มีทั้ง 2 รูปแบบ โดยเรียกว่า DNN-1 และ DNN-2 เพื่อช่วยในการประมาณช่องสัญญาณในระบบ MIMO-OFDM ด้วยสองสถานการณ์ที่แตกต่างกันของ Fading multipath ในช่องสัญญาณ ตามโมเดล TDL-

A ที่กำหนดไว้ในเครือข่าย 5G NS โดยวิธีการประมาณค่าช่องสัญญาณตาม DNN ที่เสนอได้ทำการทดสอบเทียบกับเทคนิค LS และ MMSE โดยใช้การเข้ารหัสสัญญาณแบบ QPSK ในแง่ของความผิดพลาดการประมาณช่องสัญญาณและอัตราส่วนข้อผิดพลาดบิต (BER) ดังรูปที่ 2.20



รูปที่ 2.20 ประสิทธิภาพ BER โดยเทคนิค LS, LMMSE, DNN-1 และ DNN-2 เทียบกับ SNR [18]

จากรูปที่ 2.20 หากพิจารณาด้วยสายตาอาจจะมองไม่เห็นความแตกต่างกันมากในวิธีการคำนวณ BER และเมื่อเทียบกับวิธี LMMSE นั้น ณ ที่ 15 SNR (dB) มีค่า BER ที่ต่ำกว่า DNN-1 และ DNN-2 ซึ่งเป็นวิธีการใช้ DL นั้นเอง



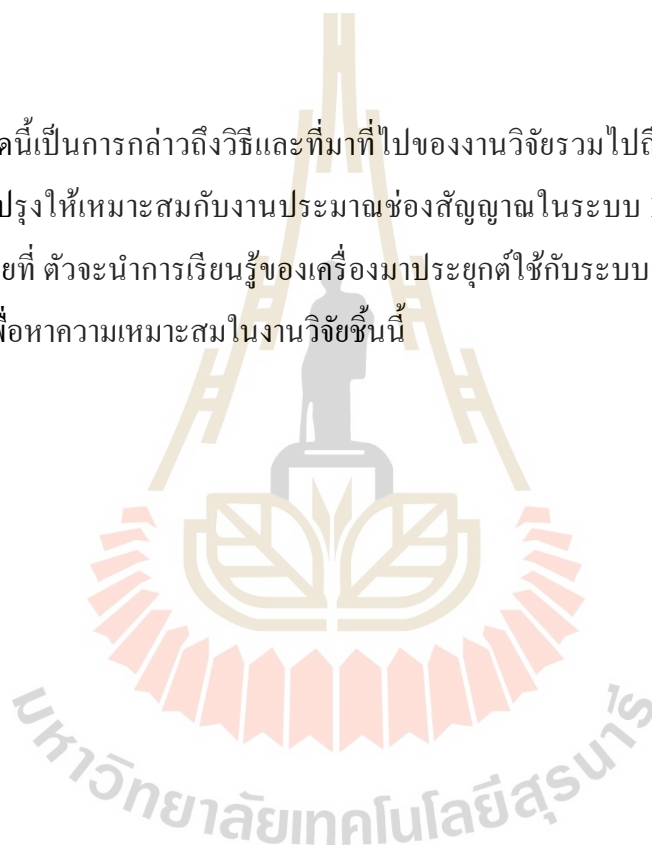
รูปที่ 2.21 ประสิทธิภาพ MSE โดยเทคนิค LS, LMMSE, DNN-1 และ DNN-2 เทียบกับ SNR [18]

จากรูปที่ 2.21 จากผลลัพธ์ที่แสดงจะเห็นได้ว่า DNN-1 และ DNN-2 ที่เป็นวิธีการแบบ DL นั้นมีประสิทธิภาพดีกว่าในด้านการหาค่าความคลาดเคลื่อนเฉลี่ย (MSE) เมื่อเทียบกับ 2 เทคนิคแบบดั้งเดิมที่นิยมใช้งานกันมาอย่างยาวนาน

อาจเพราะด้วยอัลกอริทึมได้มีการจัดการกับข้อมูลที่มีปริมาณที่พอเหมาะ โดยในงานวิจัยชิ้นนี้ไม่ได้มีกระแสดังกล่าวให้เห็นถึงการคำนวณทางเวลาเป็นเพียงการนำ DL มาประยุกต์ใช้งานเพื่อพิสูจน์ความสามารถของอัลกอริทึมที่มีผลต่อระบบสื่อสารเท่านั้น

2.7 สรุป

ทั้งหมดนี้เป็นการกล่าวถึงวิธีและที่มาที่ไปของงานวิจัยรวมไปถึงการเลือกใช้อัลกอริทึมเพื่อนำมาปรับปรุงให้เหมาะสมกับงานประมาณช่องสัญญาณในระบบ Massive MIMO โดยทางผู้วิจัยมีเป้าหมายที่ จะนำการเรียนรู้ของเครื่องมาประยุกต์ใช้กับระบบสื่อสารและการประมาณช่องสัญญาณ เพื่อหาความเหมาะสมในงานวิจัยชิ้นนี้



บทที่ 3

วิธีดำเนินการวิจัย

3.1 กล่าวนำ

ในบทนี้จะเป็นการอธิบายขั้นตอนการทำงานของวิธีวิจัยที่ใช้การประมาณช่องสัญญาณในระบบโมโฆขนานใหญ่ โดยทางผู้วิจัยได้นำเทคนิคการเรียนรู้ของเครื่องมาประยุกต์ใช้งาน ได้แก่ Extreme Learning Machine (ELM), Regularized Extreme Learning Machine (RELM) และ Outliner-Robust Extreme Learning Machine (ORELM) ซึ่งเป็นการศึกษาหาผลลัพธ์ความเป็นไปได้จากการนำเทคนิคเหล่านี้มาใช้ประมาณช่องสัญญาณ จากนั้นนำเทคนิคการประมาณช่องสัญญาณแบบเดิมคือวิธี LS และ MMSE มาเป็นตัวเปรียบเทียบ เพื่อทำการวัดประสิทธิภาพด้วยวิธีการหาความคลาดเคลื่อนเฉลี่ย (MSE), อัตราความผิดพลาดการรับส่งข้อมูล (BER) และ ความจุของช่องสัญญาณ (Channel Capacity)

3.2 การจำลองระบบ

การจำลองผลของระบบโดยใช้โปรแกรม MATLAB เป็นวิธีที่ได้รับการยอมรับในหมู่นักวิจัยทั่วไปและมีมาตรฐานในการทดสอบ ซึ่งในการทดสอบค่าความคลาดเคลื่อนในระบบโมโฆขนานใหญ่ ทางผู้วิจัยได้ทำการจำลองสายอากาศทางภาคส่ง 128 ต้น และสายอากาศทางภาครับ 128 ต้น เพื่อทำการประมาณช่องสัญญาณสามารถแสดงสมการได้ดังนี้

$$Y = H \times X + n \quad (3.1)$$

โดยที่ Y คือ สัญญาณทางภาครับ M_R

H คือ ช่องสัญญาณระหว่างภาคส่งและภาครับ $M_T \times M_R$

X คือ สัญญาณทางภาคส่ง $M_T \times 1$

n คือ สัญญาณรบกวน Add white Gaussian noise ทางภาครับ M_R

ซึ่งสามารถเขียนอยู่ในรูปแบบเมทริกซ์ได้ดังนี้

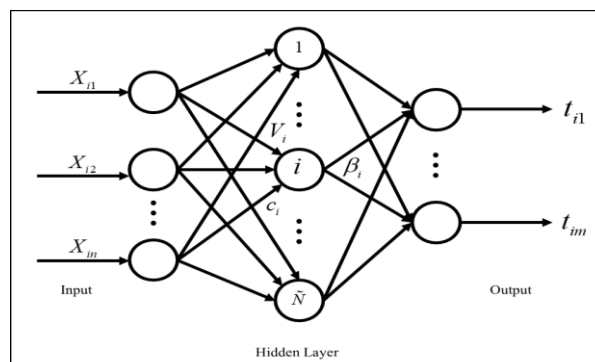
$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{M_R} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1,M_T} \\ h_{21} & h_{22} & \cdots & h_{2,M_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R,1} & h_{M_R,2} & \cdots & h_{M_R,M_T} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{M_T} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{M_R} \end{bmatrix} \quad (3.2)$$

การประมาณช่องสัญญาณ โดยการประยุกต์ใช้การเรียนรู้ของเครื่องที่ทางผู้วิจัยทำการศึกษาคือ 3 วิธี ได้แก่ Extreme Learning Machine, Regularized Extreme Learning Machine และ Outliner-Robust Extreme Learning Machine โดยมีอีก 2 วิธีทำเป็นตัวเปรียบเทียบประสิทธิภาพคือ Least Square Estimation และ Minimum Mean Square Estimation ซึ่งทั้ง 5 วิธีมีความยากง่ายและซับซ้อนต่างกัน ดังนี้

3.3 เทคนิคการเรียนรู้ของเครื่อง Extreme Learning Machine (ELM)

โดยที่เทคนิคนี้ได้มีการคิดค้นขึ้นเมื่อปี 2004 และการเรียนรู้เชิงลึก (Deep learning) เช่นเดียวกัน เนื่องจากมีประสิทธิภาพในการประมวลผลทางคอมพิวเตอร์สูง โดยมีเป้าหมายก็คือ สภาพแวดล้อมของการประมวลผลที่สูงและข้อมูลที่มีความยืดหยุ่นหลากหลาย รวมไปถึง อัลกอริทึมการเรียนรู้ที่มีประสิทธิภาพ โดยทั้ง 3 เรื่องนี้เป็นตัวขับเคลื่อนให้ การเรียนรู้ของเครื่อง นั้นมีการดำเนินมาตั้งแต่ปี 2004 [17]

จึงเป็นขั้นตอนวิธีที่มีความเร็ว มีโครงสร้างการทำงานแบบโครงข่ายประสาทแบบป้อนไปข้างหน้า Single-hidden layer feedforward neural networks (SLFNs) โครงข่ายที่มีการแบ่งเป็นชั้น ๆ ชัดเจน และข้อมูลจะถูกป้อนจากชั้นแรกแล้วไหลไปเป็นทางเดียวไปเรื่อย ๆ จนถึงชั้นสุดท้ายโดย ไม่มีการย้อนกลับไปยังชั้นก่อนหน้า



รูปที่ 3.1 โครงสร้างระบบประสาทของ Extreme Learning Machine

จากรูปภาพที่ 3.1 เราได้ทำการแสดงให้เห็นถึง โครงสร้างการทำงานของ ELM ซึ่งมีตัวแปรต่าง ๆ สามารถเขียนอยู่ในรูปแบบสมการดังนี้

$$\hat{t}_j = \sum_{i=1}^N \beta_i o(c_i \cdot X_j + V_i), j = 1, 2, \dots, N \quad (3.3)$$

โดยที่	X_i	คือ ค่าอินพุตของข้อมูล
	t_i	คือ ค่าเอาต์พุตของข้อมูล
	c_i	คือ ค่าน้ำหนักอินพุตของข้อมูลชั้นซ่อน $c_i = [c_{i1}, c_{i2}, \dots, c_{im}]^T$
	β_i	คือ ค่าน้ำหนักเอาต์พุตของข้อมูลชั้นซ่อน $\beta = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$
	V_i	คือ ค่าไบอัสของข้อมูลชั้นซ่อน
	o	คือ ค่า activation function ของข้อมูลชั้นซ่อน

จากตัวอย่างข้อมูลทั้งหมด N ขั้วต้นสามารถเขียนเทียบเคียงสมการที่ 3.1 ได้เป็นระบบเชิงเส้นคือ

$$H\beta = T \quad (3.4)$$

เมื่อ

$$H(c_1, \dots, c_N, V_1, \dots, V_N, X_1, \dots, X_N) = \begin{bmatrix} o(c_1 \cdot X_1 + V_1) & \dots & o(c_N \cdot X_1 + V_N) \\ \vdots & \dots & \vdots \\ o(c_1 \cdot X_N + V_1) & \dots & o(c_N \cdot X_N + V_N) \end{bmatrix}_{N \times N} \quad (3.5)$$

โดยที่ H คือ โหนดชั้นซ่อนเอาต์พุตเมทริกซ์ของ SLFNs ถูกแทนด้วยข้อมูลตามจำนวน i แถวและ j คอลัมน์ของผลลัพธ์จากเมทริก H ในโหนดชั้นซ่อนและมี β กับ T ดังนี้

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}_{N \times m} \quad T = \begin{bmatrix} T_1 \\ \vdots \\ T_N \end{bmatrix}_{N \times m} \quad (3.6)$$

เพื่อให้สามารถหาค่าเครือข่ายของฟังก์ชันต้นทุน (Network cost function) $\|T - H\beta\|$ ได้น้อยที่สุด โดยที่ เมื่อ T คือเมทริกซ์เป้าหมายของเอาต์พุตโดยทฤษฎีของ ELM ได้กล่าวว่า พารามิเตอร์ของ โหนดชั้นซ่อน c_i และ X_i สามารถกำหนดค่าได้จากการสุ่มโดยไม่ต้องพิจารณาข้อมูลที่เข้ามา

ดังนั้นสมการที่ 3.1 จะกลายเป็นสมการ โมเดลเชิงเส้น (Linear Model) และการหาค่า น้ำหนักของเอาต์พุตสามารถวิเคราะห์โดยหาค่าจากกำลังสองน้อยสุดของระบบเชิงเส้นหาได้ดัง สมการต่อไปนี้

$$\hat{\beta} = H^{\oplus} T \quad (3.7)$$

กำหนดให้ H^{\oplus} คือการหาเมทริกซ์ผกผันแบบมัวร์เพนโรส (Moore-Penrose) ของเมทริกซ์เอาต์พุต ชั้นซ่อน H เนื่องจากปกติแล้วจำนวน training sample จะมากกว่าจำนวนโหนดชั้นซ่อนสามารถ เขียนใหม่ได้เป็น

$$\hat{\beta} = (H^H H)^{-1} H^H T \quad (3.8)$$

3.4 เทคนิคการเรียนรู้ของเครื่อง Regularized Extreme Learning Machine (RELM)

โดยสรุปแล้วส่วนสำคัญของ ELM คือการลดข้อผิดพลาดในการฝึก แต่สิ่งนี้อาจก่อให้เกิด การความไม่เหมาะสมของผลลัพธ์ (overfitting) มากเกินไปซึ่งจะทำให้ความแม่นยำในการทำนาย ลดลง ซึ่งพิจารณาตาม Bartlett's theory [10] สำหรับ SFLNs แล้วยิ่งบรรทัดฐานของน้ำหนักและ ข้อผิดพลาดในการฝึกมีขนาดเล็กลง ประสิทธิภาพการสรุปทั่วไปที่ดีกว่าเดิมในระบบเครือข่าย ประสาทเทียม หรือกล่าวอีกนัยหนึ่งคือ โมเดล ELM ที่มีความสามารถในการวางนัยทั่วไปที่ดีควรมี การแลกเปลี่ยนข้อมูลที่ดีที่สุดในระหว่างข้อผิดพลาดในการฝึกกับบรรทัดฐานของน้ำหนักเอาต์พุต โดยที่สามารถแบ่งให้เป็นสัดส่วนได้โดยการนำพารามิเตอร์ C มาทำให้เป็นมาตรฐานดังนั้นเมื่อ เปรียบเทียบกับ ELM ธรรมดา ELM ประเภทนี้จึงเรียกว่า RELM เนื่องจากเงื่อนไขการทำให้เป็น มาตรฐานและโดยปกติจะมีประสิทธิภาพในการกำหนดลักษณะทั่วไปที่ดีกว่า

สำหรับคุณสมบัติหลักของ RELM ที่แตกต่างจาก ELM คือการลดข้อผิดพลาดในการ ฝึกอบรมพร้อมกันและการกำหนดมาตรฐานของน้ำหนักเอาต์พุตด้วยพารามิเตอร์การทำให้เป็น ระเบียบ สิ่งนี้สามารถแสดงเป็น

$$\min_{\beta} C \|T - H\beta\|_2^2 + \|\beta\|_2^2 \quad (3.9)$$

โดยที่สามารถเปลี่ยนให้อยู่ในรูปแบบอย่างง่ายเทียบเท่ากับปัญหาการเพิ่มประสิทธิภาพที่มีข้อจำกัดดังต่อไปนี้

$$\min_{\beta} C \|a\|_2^2 + \|\beta\|_2^2 \quad (3.10)$$

โดยที่กำหนดให้ $a = T - H\beta$ และปรับเปลี่ยนให้อยู่ในรูปแบบของลากรองจ์ (Lagrangian) กำหนดได้ดังนี้

$$L(\beta, a, \lambda) = C \|a\|_2^2 + \|\beta\|_2^2 + \lambda^H (T - H\beta - a) \quad (3.11)$$

โดยที่ $a = [a_1, a_2, a_3, \dots, a_N]^T$ คือตัวแปรของข้อผิดพลาดตามจำนวนของ N และค่า λ เป็นเวกเตอร์คอนจูเกตของตัวคูณแบบลากรองจ์ ตามทฤษฎีบท Karush – Kuhn – Tucker (KKT) ซึ่งเงื่อนไขการเพิ่มประสิทธิภาพนั้นกำหนดโดย

$$\begin{cases} \frac{\partial L}{\partial \beta} = 0 \rightarrow 2\beta - H^T \lambda = 0 \\ \frac{\partial L}{\partial a} = 0 \rightarrow 2Ca - \lambda = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \rightarrow T - H\beta - a = 0 \end{cases} \quad (3.12)$$

โดยกระบวนการนี้สามารถไปแก้ไขปัญหาค่า β ได้คือ

$$\hat{\beta} = (H^H H + \frac{I}{C})^{-1} H^H T \quad (3.13)$$

แต่มีสิ่งที่จะต้องระวังคือ เมื่อจำนวนของตัวฝึกอบรมน้อยกว่าจำนวนโหนดชั้นซ่อน เพื่อลดจำนวนการคำนวณลงสามารถทำได้โดยวิธี

$$\hat{\beta} = H^H (HH^H + \frac{I}{C})^{-1} T \quad (3.14)$$

3.5 เทคนิคการเรียนรู้ของเครื่อง Outlier-Robust Extreme Learning Machine (ORELM)

โดยทั่วไปค่าผิดปกติใช้เพียงส่วนน้อยของตัวอย่างการฝึกอบรมทั้งหมดสำหรับข้อผิดพลาดในการฝึกอบรม a ลักษณะนี้อาจอธิบายได้อย่างไม่ละเอียด ซึ่งที่ทราบกันดีคือความเบาบางสามารถสะท้อนได้ดีกว่าโดยใช้ l_0 -norm แทนที่จะเป็น l_2 -norm ดังนั้นเราจึงมองหาค่าน้ำหนักเอาต์พุต β แทนด้วย l_2 -norm ขนาดเล็กลง เพื่อให้ข้อผิดพลาดในการฝึก a นั้นลดลง เช่น

$$\min_{\beta} C \|a\|_0 + \|\beta\|_2^2 \quad (3.15)$$

โดยที่ขึ้นอยู่กับ $a = T - H\beta$

อย่างไรก็ตามในสมการที่ 3.15 นั้นเป็นปัญหาในการเขียน non-convex ซึ่งมีวิธีที่ง่ายกว่าในการแก้ปัญหาคือ การเขียนรูปแบบขึ้นมาใหม่ให้อยู่ในรูปแบบ convex relaxation ที่สามารถจับต้องได้โดยไม่สูญเสียรูปแบบลักษณะการกระจายจากการตรวจพบและวิเคราะห์องค์ประกอบหลักที่มีประสิทธิภาพที่สามารถเข้าถึงได้ด้วย l_1 -norm เห็นได้ชัดว่าการแทนที่ l_0 -norm ด้วย l_1 -norm ในสมการที่ 3.15 ไม่เพียงแต่พิสูจน์ความเบาบางเท่านั้น แต่นำไปสู่การลดขนาดโดยรวมลงอีกด้วย เป็นผลทำให้ได้รูปแบบดังนี้

$$\min_{\beta} \|a\|_1 + \frac{1}{C} \|\beta\|_2^2 \quad (3.16)$$

โดยที่ขึ้นอยู่กับ $a = T - H\beta$

ในสมการที่ 3.16 เป็นปัญหาการเพิ่มประสิทธิภาพ convex ที่มีข้อจำกัดและเหมาะสมอย่างยิ่งกับช่วงที่ใช้งานได้ของวิธีการ augmented Lagrange multiplier (ALM) ในทำนองเดียวกันฟังก์ชันแบบลากรองจ์ ที่เพิ่มขึ้นดังนี้

$$L_{\mu}(a, \beta, \lambda) = \|a\|_1 + \frac{1}{C} \|\beta\|_2^2 + \lambda^H (T - H\beta - a) + \frac{\mu}{2} \|(T - H\beta - a)\|_2^2 \quad (3.17)$$

กำหนดให้ λ เป็นเวกเตอร์คอลลัมน์ของตัวคูณแบบลากรองจ์ และ μ คือฟังก์ชันพินอลตีตามคำแนะนำของ [14] โดยที่เราจะเลือกให้ $\mu = 2N / \|T\|_1$ โดยในอัลกอริทึม ALM จะประมาณค่า (a, β) และตัวคูณฟังก์ชันแบบลากรองจ์ λ ที่เหมาะสมที่สุด โดยการย่อค่าที่ฟังก์ชันแบบลากรองจ์ลงเรื่อย ๆ ในระยะเวลาที่กำหนดจะได้ (λ_k, μ) สำหรับรูปแบบการทำซ้ำของ ALM คือ

$$\begin{cases} (a_{k+1}, \beta_{k+1}) = \arg \min_{a, \beta} L_\mu(a, \beta, \lambda_k) \\ \lambda_{k+1} = \lambda_k + \mu(T - H\beta_{k+1} - a_{k+1}) \end{cases} \quad (3.18)$$

การย่อขนาดในสมการแรกของสมการที่ 3.16 ของการทำซ้ำในอัลกอริทึม ALM สามารถทำได้โดยใช้เทคนิคทิศทางกราดลับที่ใช้อย่างแพร่หลายซึ่งจะอัปเดตค่า a และ β แบบอนุกรมทั้งสองตัวที่ไม่รู้จัก ให้มีความแม่นยำมากขึ้นในอัลกอริทึม ALM โดยมีเป้าหมายที่จะแก้ปัญหาต่อไปนี้เพื่อสร้างการทำซ้ำใหม่ดังนี้

$$\begin{cases} \beta_{k+1} = \arg \min_{\beta} L_\mu(a, \beta, \lambda_k) \\ a_{k+1} = \arg \min_a L_\mu(a, \beta_{k+1}, \lambda_k) \\ \lambda_{k+1} = \lambda_k + \mu(T - H\beta_{k+1} - a_{k+1}) \end{cases} \quad (3.20)$$

โดยที่สมการนี้อ้างอิงจาก [15] ซึ่ง β_{k+1} มียอมรับวิธีแก้ปัญหาที่ชัดเจนดังต่อไปนี้

$$\beta_{k+1} = (H^H H + 2/C\mu I)^{-1} H^H (T - a_k + \lambda_k / \mu) \quad (3.20)$$

โดยที่ a_{k+1} แสดงดังนี้

$$a_{k+1} = \text{shrink}(T - H\beta_{k+1} + \lambda_k / \mu, 1/\mu) \quad (3.21)$$

$$\max\{T - H\beta_{k+1} + \lambda_k / \mu, 0\} \times \text{sign}(T - H\beta_{k+1} + \lambda_k / \mu) \quad (3.22)$$

โดยที่ \times แสดงถึงการคูณตามองค์ประกอบอย่างชาญฉลาด (the element-wise multiplication.)

จะเห็นได้ชัดเจนว่าจำนวนการคำนวณของการวนซ้ำแต่ละครั้งส่วนใหญ่ใช้ในการแก้ค่าที่ผกผันของ $H^H H + 2/C\mu I$ โดยในการวนซ้ำแต่ละครั้งจะมีเมทริกซ์เดียวกัน $(H^H H + 2/C\mu I)^{-1} H^H$ ซึ่งสามารถคำนวณค่าล่วงหน้าได้ก่อนที่จะเริ่มการวนซ้ำทั้งหมด จึงทำให้การคำนวณมีความเร็วมาก ในขณะที่เดียวกัน มีข้อสังเกตว่าเทคนิคความต่อเนื่องที่แตกต่างกัน μ สามารถเร่งการบรรจบกันด้วยการวนซ้ำน้อยลง [11-12] แต่อย่างไรก็ตามสมการนี้ต้องใช้การคำนวณต่าง ๆ ของ $(H^H H + 2/C\mu I)^{-1} H^H$ ซึ่งจะนำไปสู่การคำนวณที่มีจำนวนมากขึ้น แม้ว่าจะสามารถใช้เทคนิค SVD ที่รวดเร็วกว่าก็ตาม [13]

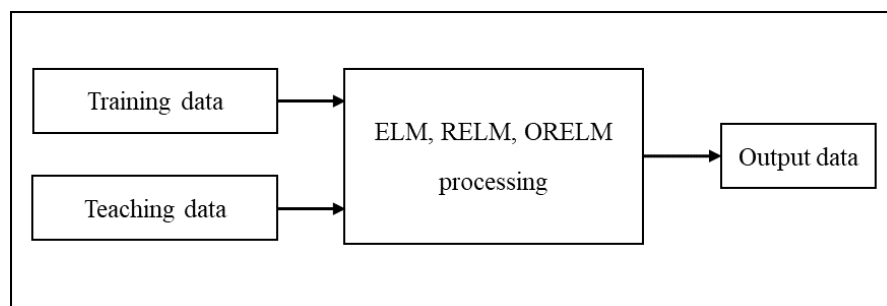
นอกจากนี้ควรสังเกตว่า แม้ว่าวิธีการทำซ้ำนี้ไม่จำเป็นต้องเร็วที่สุดในการเข้าถึงความแม่นยำที่สูงมากเมื่อ training data ไม่มีค่าผิดปกติ แต่ก็ยังเป็นข้อมูลที่มีประสิทธิภาพมากที่สุดในการได้รับระดับความแม่นยำที่ดีที่สุด

แต่เมื่อใดก็ตามที่ข้อมูลมีค่าผิดปกติ ซึ่งเป็นกรณีที่เกี่ยวข้องกับการใช้งานจริงมากที่สุด เพื่อความสะดวก จึงจำเป็นต้องมีการหยุดการทำซ้ำเมื่อจำนวนการวนซ้ำถึงขีดจำกัด (MaxIter) ที่กำหนดไว้ล่วงหน้า

3.6 การนำเทคนิคการเรียนรู้ของเครื่องมาประยุกต์หาการประมาณช่องสัญญาณ

จากนี้จะเป็นการประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องจากหัวข้อที่ 3.2 ถึง 3.5 โดยที่แสดงดังรูปต่อไปนี้

ซึ่งทางผู้วิจัยจะกำหนดให้ ข้อมูลนำเข้า (input) หรือ Training data เป็นช่องสัญญาณที่ได้รับได้ก็คือพารามิเตอร์ Y และ H คือข้อมูลนำออก (output) หรือ Teaching data มีค่าเป็นช่องสัญญาณระหว่าง พารามิเตอร์ Y และ X ในระบบสื่อสารของหัวข้อที่ 3.2 โดยที่แสดงโครงสร้างดังรูปที่ 3.2



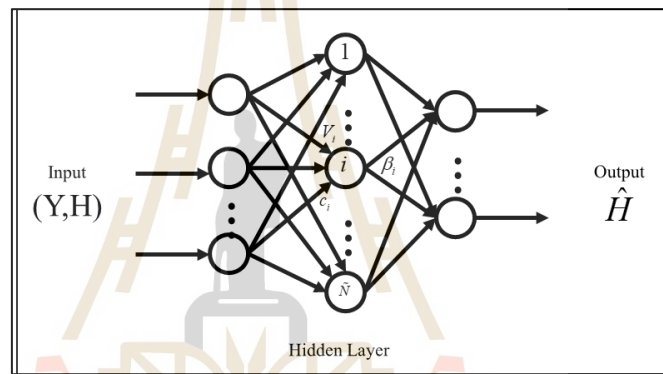
รูปที่ 3.2 โครงสร้างการทำงานในการเตรียมข้อมูล

โดยที่ช่องสัญญาณ H เป็นการสุ่มค่าขึ้นมาจากการจำลองโปรแกรม MATLAB ให้มีขนาดเท่ากับจำนวนสายอากาศส่งและภาครับที่เท่า ๆ กัน



$$H = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,M_T} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,M_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R,1} & h_{M_R,2} & \cdots & h_{M_R,M_T} \end{bmatrix} \quad (3.23)$$

และสำหรับ X นั้นคือการกำหนดสัญญาณทางภาคส่งเพราะจะมีผลโดยตรงกับสัญญาณทางภาครับ Y ดังสมการที่ 3.1 ในการเข้ารหัสสัญญาณ (Modulation) และถอดรหัสสัญญาณ (Demodulation) โดยทางผู้วิจัยได้พิจารณาการเข้ารหัสแบบพื้นฐานเพื่อที่จะมุ่งเน้นในส่วนของเทคนิคการประมาณช่องสัญญาณโดยเฉพาะ



รูปที่ 3.3 โครงสร้างการระบุข้อมูลที่ใช้ในกระบวนการเรียนรู้ของเครื่อง

โดยสิ่งที่ต้องปรับเปลี่ยนให้เหมาะสมคือจำนวนโหนดชั้นซ่อน และพารามิเตอร์ต่าง ๆ ในกระบวนการเรียนรู้ของเครื่องในแต่ละอัลกอริทึม โดยที่ RELM จะมีพารามิเตอร์บางตัวที่คอยควบคุมความผันผวนและลดการ Overfitting ของข้อมูลลง ทำให้สามารถได้ข้อมูลขาออกที่มีความเสถียรมากขึ้นและลดระยะเวลาในการทำงานลงบางช่วง พารามิเตอร์ตัวนั้นถูกเรียกว่า The regularization parameter ; C นั้นเอง โดยที่การปรับพารามิเตอร์ต่าง ๆ นั้น ไม่ได้มีข้อมูลที่สามารถอธิบายได้ดีนัก แต่ขึ้นอยู่กับผลลัพธ์ของข้อมูลที่ผ่านมาแล้ว เช่น ระยะเวลาที่ทำงานรวมไปถึงประสิทธิภาพหลังผ่านกระบวนการเรียนรู้ของเครื่อง ดังนั้นข้อมูลขาออก (output) ที่จะได้คือ ค่าประมาณช่องสัญญาณที่ผ่านกระบวนการของการเรียนรู้ของเครื่องดังนี้

ค่าประมาณช่องสัญญาณแบบ ELM

$$\hat{H}_{ELM} \quad (3.24)$$

ค่าประมาณช่องสัญญาณแบบ RELM

$$\hat{H}_{RELM} \quad (3.25)$$

ค่าประมาณช่องสัญญาณแบบ ORELM

$$\hat{H}_{ORELM} \quad (3.26)$$

จากนั้นผู้วิจัยจะนำข้อมูลที่ได้จากแต่ละอัลกอริทึมไปคำนวณหาประสิทธิภาพต่าง ๆ เพื่อบรรลุถึงความเป็นไปได้ต่อไป

3.7 การนำเทคนิคการประมาณช่องสัญญาณ LS และ MMSE มาใช้งาน

หลังจากที่ทำการประมาณช่องสัญญาณที่ใช้เทคนิคต่าง ๆ มาประยุกต์แล้ว จากนั้นทางผู้วิจัยจึงต้องนำเสนอการประมาณช่องสัญญาณแบบต่าง ๆ เพื่อเอามาใช้เป็นตัวเปรียบเทียบประสิทธิภาพในด้านอื่น ๆ โดยเราใช้อยู่ 2 วิธีคือ LS และ MMSE ซึ่งมีสมการดังต่อไปนี้

$$\hat{H}_{(N)LS} = YX^H (XX^H)^{-1} \quad (3.27)$$

$$\hat{H}_{(N)MMSE} = YX^H \left(XX^H + \frac{\sigma_n^2}{\sigma_h^2} I \right)^{-1} \quad (3.28)$$

โดยจะนำเทคนิคที่กล่าวถึงในบทที่ 2 เกี่ยวข้องกับทฤษฎีของการประมาณช่องสัญญาณแบบ LS และ MMSE โดยสรุปแล้วทั้ง 2 เทคนิคการประมาณช่องสัญญาณแบบนี้ได้ถูกนำไปใช้งานกันอย่างแพร่หลายและอีกทั้งยังเป็นตัววัดประสิทธิภาพในการประมาณช่องสัญญาณในเทคนิคต่าง ๆ ที่นักวิจัยได้คิดค้นขึ้นมา เนื่องจากด้วยวิธีคิดนั้นไม่มีความซับซ้อนมากนักนั่นเอง

3.8 วิธีการเทียบประสิทธิภาพแบบ MSE

ในการวัดประสิทธิภาพโดยทั่วไปแล้วมีหลายวิธีมากที่สามารถบ่งบอกถึงประสิทธิภาพโดยรวม แต่วิธีที่เป็นมาตรฐานหรือที่นิยมใช้งานกันก็คือ การวัดค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

(Mean Square Error) ซึ่งเป็นที่นิยมใช้ในงานประเภท Regression โดยมีขั้นตอนในการวัดประสิทธิภาพดังนี้

$$MSE_N = \sum_{i=1}^N \|H - \hat{H}\|^2 \tag{3.29}$$

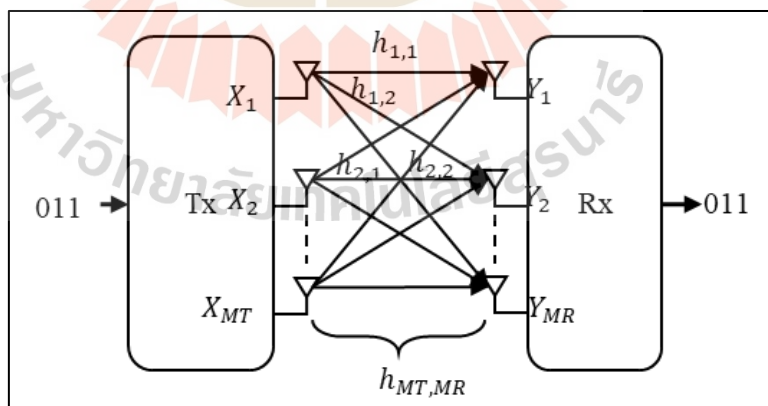
โดยที่ H คือ ช่องสัญญาณจริงจากการจำลองระบบสื่อสาร

\hat{H} คือ ช่องสัญญาณที่ประมาณได้โดยวิธีหรือเทคนิคต่าง ๆ

โดยที่ \hat{H} คือ วิธีดังนี้ LS, MMSE, ELM, RELM, ORELM เป็นต้น ซึ่งจะพิจารณาจำนวนค่าความคลาดเคลื่อนเฉลี่ยกับค่า SNR ในแต่ละช่วง นั้นจากวิธี MSE นั้นเป็นการยกกำลังสองเฉลี่ย โดยไม่คำนึงถึงค่าที่ติดลบจึงสามารถระบุได้ว่าช่วงของข้อมูลไหนมีความคลาดเคลื่อนต่างไปจากเดิมมากน้อยเพียงใด

3.9 วิธีการเทียบประสิทธิภาพแบบ BER

จากที่ได้วัดประสิทธิภาพในเรื่องค่าความคลาดเคลื่อนไปที่เรียบร้อย ทางผู้วิจัยจึงอยากทราบถึงประสิทธิภาพในการรับส่งข้อมูลว่าการประมาณช่องสัญญาณที่รับได้ในแต่ละวิธีนั้นมีอัตราความผิดพลาดของข้อมูลมากน้อยเพียงใด สิ่งนี้จะเป็นอีกวิธีที่บ่งบอกได้ดังรูปภาพต่อไปนี้



รูปที่ 3.4 ลักษณะการส่งข้อมูลเข้ารหัสและถอดรหัสแบบง่าย

ซึ่งวิธีหาอัตราข้อผิดพลาด มีวิธีการหาแบบการกลับสมการที่ (3.1) โดย $X = Y / H$ ซึ่งเป็น การกลับสมการข้อมูลของภาคส่ง โดยที่เราจะนำ X ที่หาได้มาทำการหาอัตราผิดพลาดของข้อมูล

โดยมีการใช้คำสั่งหรือฟังก์ชันใน MATLAB เรียกว่า qammod และ qamdemod ซึ่งเป็นคำสั่งในการเข้ารหัสและถอดรหัสข้อมูล

3.10 วิธีการเทียบประสิทธิภาพแบบการหาความจุช่องสัญญาณ

โดยวิธีการเทียบประสิทธิภาพความจุช่องสัญญาณนั้นได้อ้างอิงจากสมการที่ (2.26) ซึ่งเป็นการกล่าวถึงความจุช่องสัญญาณในระบบโม โม โดยเปลี่ยนช่องสัญญาณจากเดิม H ให้กลายเป็นช่องสัญญาณที่ประมาณได้ในแต่ละเทคนิคของการประมาณช่องสัญญาณดังมีดังนี้

สำหรับความจุช่องสัญญาณโดยการประมาณช่องสัญญาณแบบ LS แสดงสมการดังต่อไปนี้

$$C = \log_2 \det \left[I_{M_R} + \frac{P_t}{P_n M_T} H_{LS} H_{LS}^H \right] \quad (3.30)$$

สำหรับความจุช่องสัญญาณโดยการประมาณช่องสัญญาณแบบ MMSE แสดงสมการดังต่อไปนี้

$$C = \log_2 \det \left[I_{M_R} + \frac{P_t}{P_n M_T} H_{MMSE} H_{MMSE}^H \right] \quad (3.31)$$

สำหรับความจุช่องสัญญาณโดยเทคนิคประยุกต์ในการประมาณช่องสัญญาณแบบ ELM แสดงสมการดังต่อไปนี้

$$C = \log_2 \det \left[I_{M_R} + \frac{P_t}{P_n M_T} H_{ELM} H_{ELM}^H \right] \quad (3.32)$$

สำหรับความจุช่องสัญญาณโดยเทคนิคประยุกต์ในการประมาณช่องสัญญาณแบบ RELM แสดงสมการดังต่อไปนี้

$$C = \log_2 \det \left[I_{M_R} + \frac{P_t}{P_n M_T} H_{RELM} H_{RELM}^H \right] \quad (3.33)$$

สำหรับความจุช่องสัญญาณ โดยเทคนิคประยุกต์ในการประมาณช่องสัญญาณแบบ ORELM แสดงสมการดังต่อไปนี้

$$C = \log_2 \det \left[I_{M_R} + \frac{P_t}{P_n M_T} H_{ORELM} H_{ORELM}^H \right] \quad (3.34)$$

3.11 สรุป

ในเนื้อหาบทนี้กล่าวถึงการสร้างแบบจำลองของระบบการสื่อสาร Massive MIMO เพื่อใช้ในการคำนวณในการประมาณช่องสัญญาณ ซึ่งยังคงความเป็นทฤษฎีบทของระบบการสื่อสารและการยกสมมุติฐานว่าเมื่อจำนวนสายอากาศเพิ่มขึ้นจะมีผลดีหรือผลเสียต่อการประมาณช่องสัญญาณในรูปแบบไหนบ้าง โดยนำการเรียนรู้ของเครื่องมาประยุกต์ใช้งานเพราะด้วยเป็นสิ่งใหม่สำหรับในงานหลากหลายด้าน อีกทั้งยังหาความเหมาะสมของการนำมาใช้งาน โดยจะทำการวัดประสิทธิภาพต่าง ๆ เช่น การหาความจุช่องสัญญาณ อัตราความผิดพลาด และค่าความคลาดเคลื่อนเฉลี่ย เป็นต้น



บทที่ 4

ผลการดำเนินงาน

4.1 กล่าวนำ

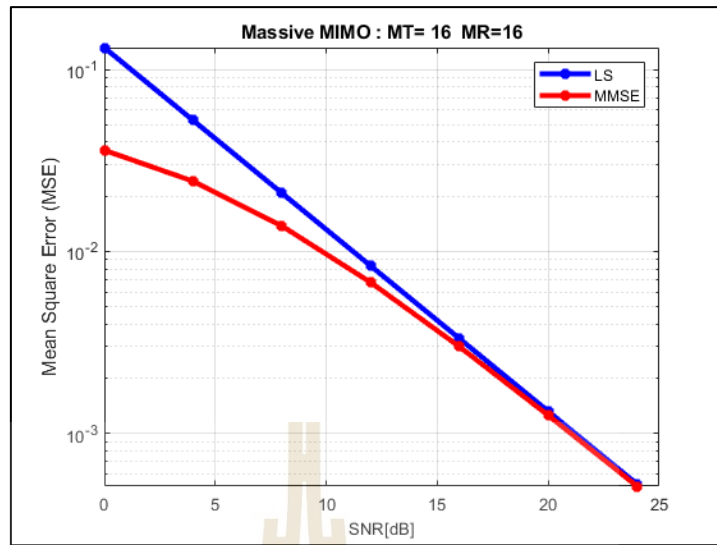
เนื้อหาก่อนหน้านี้เป็นการอธิบายถึงทฤษฎีการประมาณช่องสัญญาณในการสื่อสารระบบโมโฆขนาดใหญ่ โดยมีทฤษฎีการประมาณช่องสัญญาณแบบเดิม Least Square Error (LS) และ Minimum Mean Square Error (MMSE) และ ทฤษฎีการเรียนรู้ของเครื่อง 3 วิธีที่ผู้วิจัยนำมาใช้ เเคราะห์ผลเพื่อเปรียบเทียบ ได้แก่ Extreme Learning Machine (ELM), Regularized Extreme Learning Machine (RELM) และ Outliner-Robust Extreme Learning Machine (ORELM) ในบทนี้จะกล่าวถึงการจำลองผลที่ได้จากการใช้แต่ละเทคนิคในการเปรียบเทียบ เช่น ความคลาดเคลื่อนของ อัลกอริทึม อัตราข้อผิดพลาดในการรับส่งข้อมูล ความจุของช่องสัญญาณ และระยะเวลาประมวลผล ที่ได้โดยใช้โปรแกรม MATLAB

4.2 การทดสอบหาค่าความคลาดเคลื่อนของอัลกอริทึม

ในการทดสอบพิจารณาเปรียบเทียบค่าความคลาดเคลื่อนในระบบโมโฆขนาดใหญ่ ทางผู้วิจัยได้ทำการจำลองสายอากาศส่ง 128 ต้น และสายอากาศรับ 128 ต้น เพื่อทำการประมาณช่องสัญญาณ โดยจะใช้วิธีการดังต่อไปนี้

4.2.1 เทคนิคการประมาณช่องสัญญาณด้วยวิธี LS และ MMSE

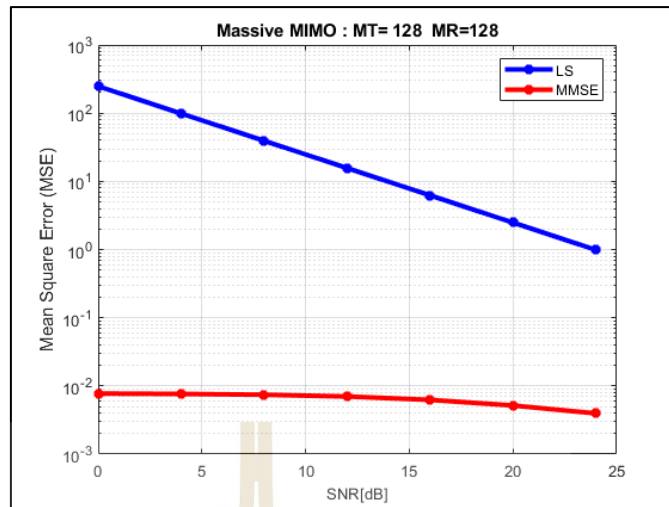
วิธีการประมาณช่องสัญญาณแบบนี้ มีความซับซ้อนของการคำนวณทางคอมพิวเตอร์น้อย ซึ่งอาจจะทำให้ผลรับคลาดเคลื่อนไปมากหากเจอข้อมูลที่มหาศาลเพราะทฤษฎี LS เป็นการคำนวณระหว่างสัญญาณที่รับได้และสัญญาณที่ถูกส่งออกมา ซึ่งไม่คำนึงถึงในส่วนของสัญญาณรบกวน ส่วนวิธีการ MMSE เป็นการมีการคำนวณทางคณิตศาสตร์เพิ่มขึ้นรวมถึงการนำสัญญาณรบกวน ทางผู้วิจัยได้กำหนดสายอากาศส่งและรับจำนวน 16, 64 และ 128 ต้น



รูปที่ 4.1 ผลทดสอบ MSE ในระบบสายอากาศส่ง 16 ต้น และสายอากาศรับ 16 ต้น



รูปที่ 4.2 ผลทดสอบ MSE ในระบบสายอากาศส่ง 64 ต้น และสายอากาศรับ 64 ต้น



รูปที่ 4.3 ผลทดสอบ MSE ในระบบสายอากาศส่ง 128 ต้น และสายอากาศรับ 128 ต้น

จากรูปที่ 4.1 ถึง 4.3 แสดงให้เห็นถึงค่าความคลาดเคลื่อนของข้อมูลในจำนวนสายอากาศที่ต่างกันคือ 16, 64 และ 128 ต้น ซึ่งสอดคล้องกับทฤษฎี ที่ยิ่งเพิ่มจำนวนสายอากาศ ค่าความคลาดเคลื่อนยิ่งสูงมากขึ้นอย่างเห็นได้ชัด ทำให้ผู้วิจัยใช้เทคนิคการประมาณช่องสัญญาณแบบใหม่คือ ELM เข้ามาวิเคราะห์หาค่าความคลาดเคลื่อนของข้อมูล รองรับการใช้งานในอนาคตที่ต้องมีการใช้ระบบโมโฆขนาดใหญ่

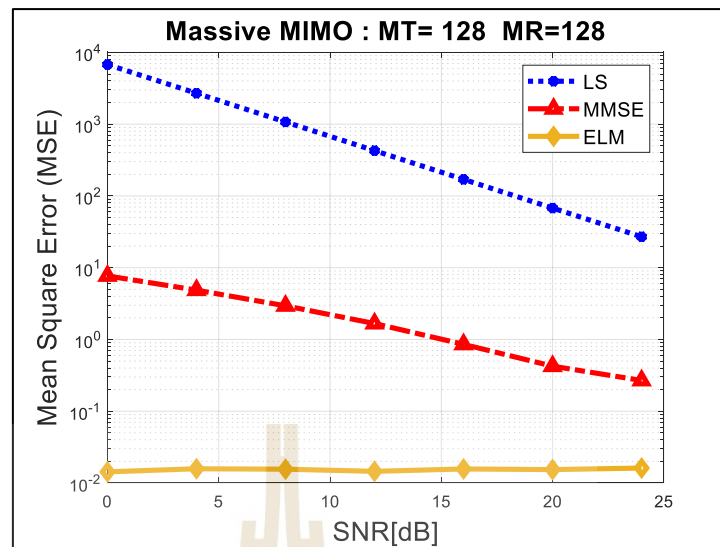
4.2.2 เทคนิคการประมาณช่องสัญญาณแบบ ELM

จากการประยุกต์ใช้วิธีการเรียนรู้ของเครื่องแบบ ELM กับ การประมาณช่องสัญญาณในระบบโมโฆขนาดใหญ่ เมื่อเทียบกับเทคนิคแบบเดิม (LS และ MMSE) ผลลัพธ์ที่ได้มีแนวโน้มที่ดีกว่าอย่างเห็นได้ชัด เพราะมีการจัดการของข้อมูลนำเข้าและข้อมูลนำออกให้สอดคล้องกับระบบสื่อสารและจำนวนสายอากาศที่ทำการทดสอบ แสดงดังรูปที่ 4.4

โดยที่ กราฟเส้นประสีน้ำเงิน แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ LS

กราฟเส้นประสีแดง แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ MMSE

กราฟเส้นทึบสีเหลือง แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ ELM



รูปที่ 4.4 การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ ELM, LS และ MMSE

จากรูปที่ 4.4 เมื่อเปรียบเทียบค่าความคลาดเคลื่อน ระหว่าง LS, MMSE และ ELM โดยใช้สายอากาศส่งและรับ ภาคละ 128 ต้น ผลที่เกิดขึ้นจะเห็นได้ว่า LS มีประสิทธิภาพต่ำสุด เนื่องจากมีค่าความคลาดเคลื่อนสูงสุด และรองลงมาเป็นเทคนิคแบบ MMSE และเห็นได้ว่าวิธี ELM มีค่าความคลาดเคลื่อนน้อยที่สุด

4.2.3 เทคนิคการประมาณช่องสัญญาณแบบ RELM

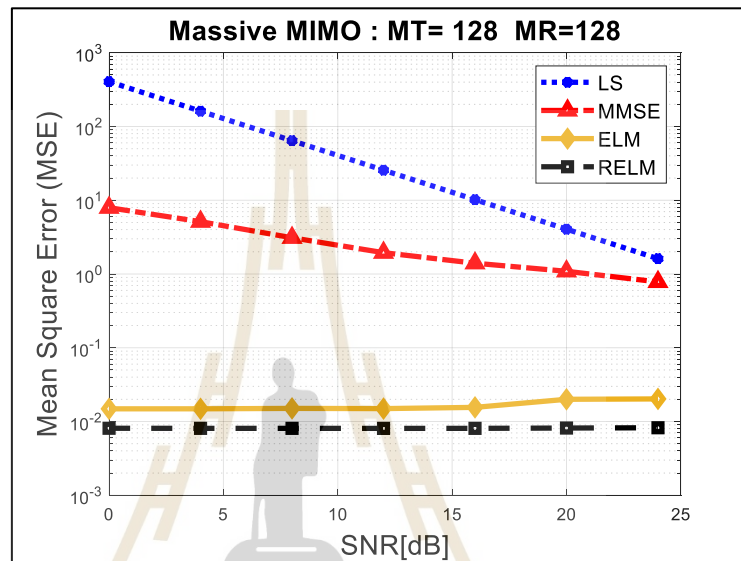
จากการประยุกต์ใช้วิธีการเรียนรู้ของเครื่องแบบ RELM กับการประมาณช่องสัญญาณในระบบ MIMO ขนาดใหญ่ เมื่อเทียบกับเทคนิคการประมาณช่องสัญญาณแบบ LS, MMSE และ ELM ผลลัพธ์ที่ได้ มีแนวโน้มดีที่สุด มีอัตราความคลาดเคลื่อนที่ต่ำลงจาก ELM เพราะมีการจัดการข้อมูลนำเข้าและข้อมูลนำออก รวมไปถึงการปรับพารามิเตอร์ที่ใช้ลดค่าความผันผวนของข้อมูล ทำให้ค่าความคลาดเคลื่อนที่เกิดขึ้นมีค่าต่ำสุด แสดงดังรูปที่ 4.5

โดยที่ กราฟเส้นประสีน้ำเงิน แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ LS

กราฟเส้นประสีแดง แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ MMSE

กราฟเส้นทึบสีเหลือง แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณ
ช่องสัญญาณแบบ ELM

กราฟเส้นทึบสีดำ แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณ
ช่องสัญญาณแบบ RELM



รูปที่ 4.5 การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ LS, MMSE, ELM และ RELM

4.2.4 เทคนิคการประมาณช่องสัญญาณแบบ ORELM

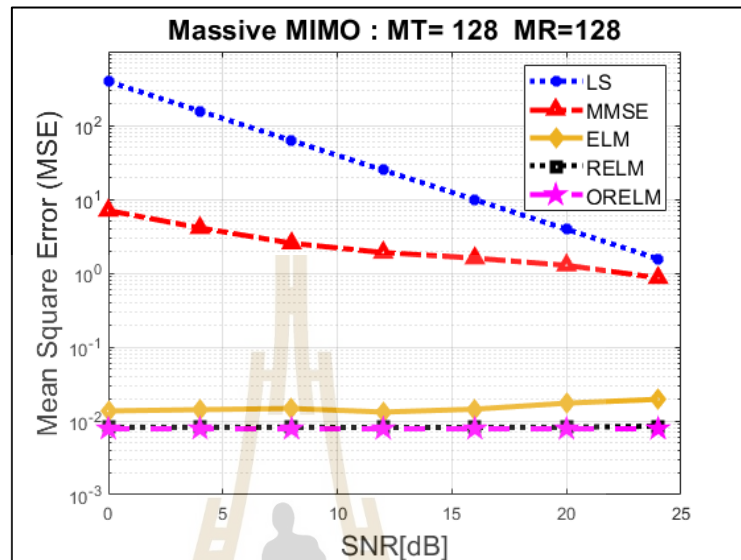
จากการประยุกต์ใช้วิธีการเรียนรู้ของเครื่อง แบบ ORELM กับการประมาณ
ช่องสัญญาณในระบบสื่อสารขนาดใหญ่ เมื่อเทียบกับเทคนิคการประมาณช่องสัญญาณแบบ LS,
MMSE, ELM และ RELM ผลลัพธ์ที่ได้นั้นมีแนวโน้ม แสดงดังรูปที่ 4.6

โดยที่ กราฟเส้นประสีแดง แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณ
ช่องสัญญาณแบบ MMSE

กราฟเส้นทึบสีเหลือง แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณ
ช่องสัญญาณแบบ ELM

กราฟเส้นทึบสีดำ แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณ
ช่องสัญญาณแบบ RELM

กราฟเส้นประสีชมพู แสดงถึงการคำนวณหาค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ ORELM

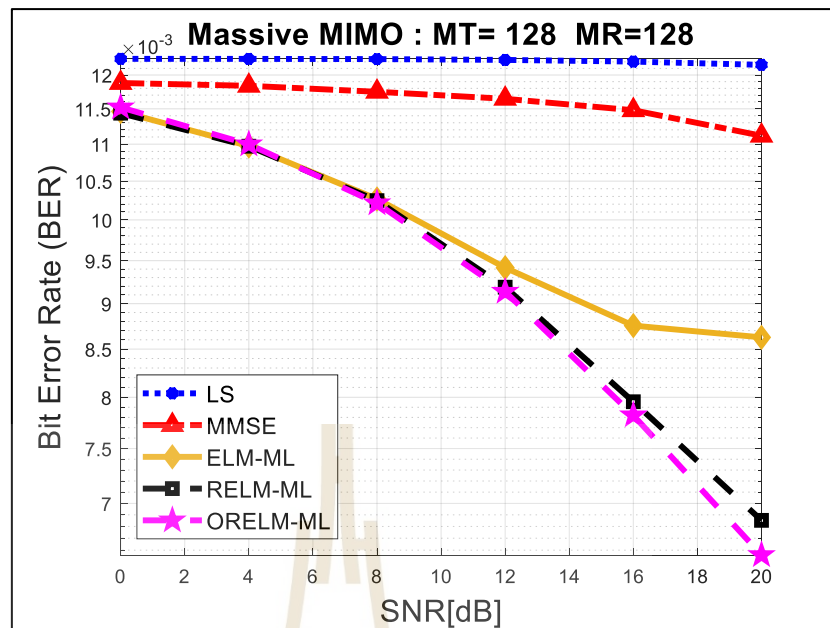


รูปที่ 4.6 การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยของการประมาณช่องสัญญาณแบบ LS, MMSE, ELM, RELM และ ORELM

จากรูปที่ 4.6 แสดงค่าความคลาดเคลื่อน ORELM มีค่าความคลาดเคลื่อนที่ต่ำสุด เพราะในตัวอัลกอริทึมมีการปรับปรุงการคำนวณที่ต้องการความแม่นยำเพิ่มขึ้น และมีการจัดการข้อมูลนำเข้าและข้อมูลนำออกรวมไปถึงการปรับพารามิเตอร์ที่ใช้ลดค่าความผันผวนของข้อมูลแต่อาจใช้เวลาในการประมวลผลเพิ่มขึ้น

4.3 การทดสอบหาอัตราบิดผิดพลาดในการส่งข้อมูล

ในการทดสอบหาอัตราบิดผิดพลาดที่เกิดขึ้นที่ภาครับต่อจำนวนบิตที่ถูกส่งมาทั้งหมดภายในช่วงเวลาหนึ่ง เพื่อพิสูจน์โดยใช้เทคนิคการประมาณช่องสัญญาณ ทั้ง 5 คือ LS, MMSE, ELM, RELM และ ORELM สามารถหาอัตราบิดผิดพลาดได้ดังรูปที่ 4.7

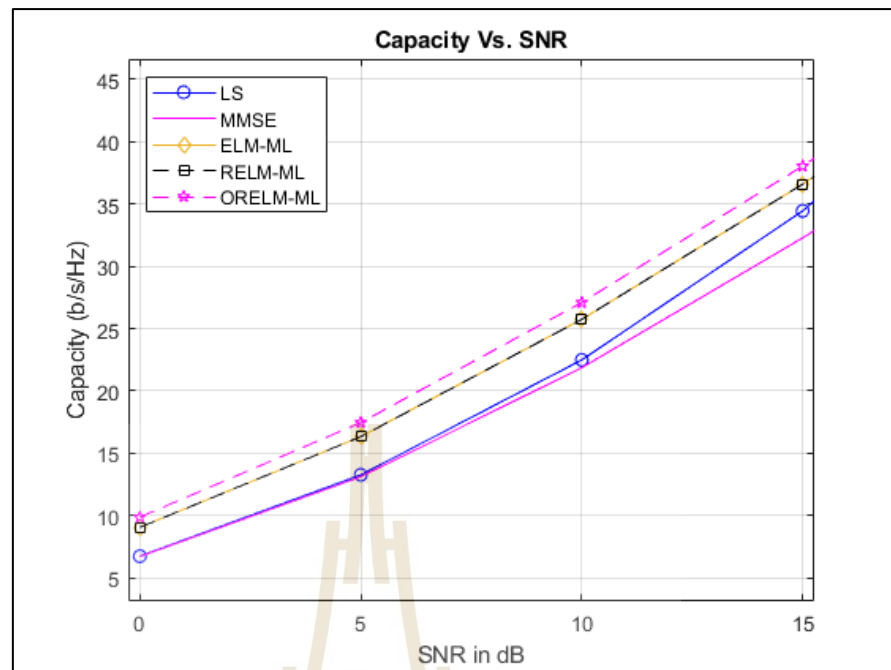


รูปที่ 4.7 การทดสอบหา BER เทียบกับ SNR

จากรูปที่ 4.7 เมื่อพิจารณาที่สายอากาศภาคส่ง 128 ต้น และสายอากาศภาครับ 128 ต้น อัตราในการรับส่งข้อมูลวิเคราะห์ได้ว่า เมื่อใช้การประมาณช่องสัญญาณแบบ LS และ MMSE มีอัตราความผิดพลาดในการรับส่งข้อมูลสูง ในขณะที่เดียวกันการประมาณช่องสัญญาณที่ใช้วิธีการเรียนรู้ของเครื่อง มีแนวโน้มที่ดีกว่าอย่างชัดเจน จากผลที่เกิดขึ้นอัตราส่งข้อมูลที่ผิดพลาดน้อยที่สุดคือวิธี ORELML รองลงมา RELM และ ELM ตามลำดับ

4.4 การทดสอบหาความจุช่องสัญญาณ

สำหรับการวัดประสิทธิภาพอีกรูปแบบคือการหาความจุของช่องสัญญาณ โดยใช้สมการ (3.30-3.34) ในการประมวลผลเปรียบเทียบ ซึ่งจะแสดงถึงอัตราเร็วในการรับส่งข้อมูล ภายใต้ช่องสัญญาณที่เกิดขึ้นแสดงดังรูปที่ 4.8



รูปที่ 4.8 ความจุช่องสัญญาณทั้ง 5 เทคนิคเทียบกับอัตราส่วนสัญญาณต่อสัญญาณรบกวน

จากรูปที่ 4.8 ผลทดสอบหาค่าความจุช่องสัญญาณเทียบกับอัตราส่วนสัญญาณต่อสัญญาณรบกวน แสดงให้เห็นว่าที่ การประมาณช่องสัญญาณแบบเดิม (LS และ MMSE) ให้ค่าความจุช่องสัญญาณที่น้อยกว่าวิธีการ ELM, RELM และ ORELM โดยที่ช่องสัญญาณและ ORELM ให้ค่าความจุของช่องสัญญาณมากที่สุด

4.5 ผลการทดสอบการประมวลผลทางเวลา (Processing Times)

การทดสอบเวลาในการประมวลผลทางคอมพิวเตอร์ ทางผู้วิจัยได้แบ่งตารางในการทดสอบออกเป็นกลุ่มเพื่อบ่งบอกถึงความชัดเจนของระยะเวลาที่ใช้ในการคำนวณของระบบคอมพิวเตอร์ดังนี้

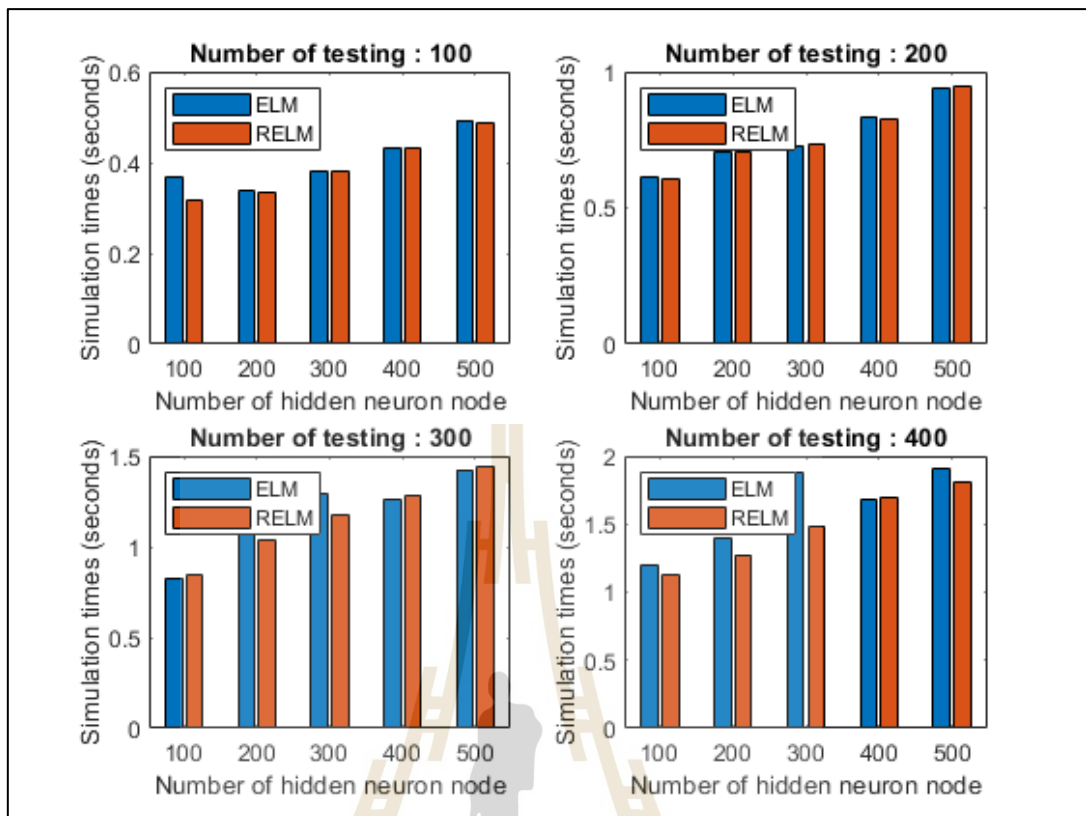
ตารางที่ 4.1 ทดสอบระยะเวลาในการคำนวณค่าความคลาดเคลื่อน (MSE)

สิ่งที่ทดสอบ	Algorithm								
	ELM			RELM			ORELM		
ค่าควบคุม ความเบี่ยงเบน: C	-			2^{30}			2^{30}		
No. Hidden layer	100	200	300	100	200	300	100	200	300
จำนวนรอบ	100 รอบ			100 รอบ			100 รอบ		
เวลาที่ทดสอบ (วินาที)	0.385	0.476	0.553	0.377	0.473	0.548	10.555	14.239	18.459

ตารางที่ 4.2 ทดสอบระยะเวลาในการคำนวณอัตราความผิดพลาดข้อมูล (BER)

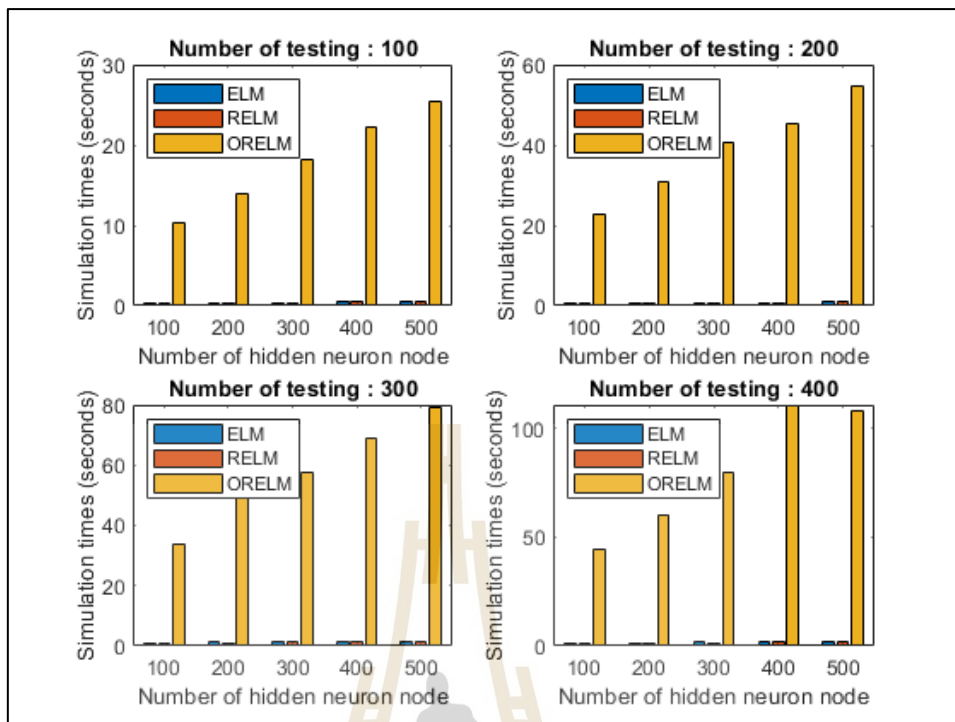
สิ่งที่ทดสอบ	Algorithm								
	ELM			RELM			ORELM		
Control : C	-			2^{30}			2^{30}		
No. Hidden layer	100	200	300	100	200	300	100	200	300
จำนวนรอบ	100 รอบ			100 รอบ			100 รอบ		
เวลาที่ทดสอบ (วินาที)	0.385	0.476	0.566	0.377	0.473	0.548	10.555	14.239	18.647

จากตารางที่ 4.1 และ 4.2 แสดงถึงการนำตัวเลขในการประมวลผลมาทดสอบหาค่าการประมวลผลทางเวลาทั้งการทดสอบค่าความคลาดเคลื่อน (MSE) และอัตราความผิดพลาดข้อมูล (BER) ผลที่เกิดขึ้นแสดงให้เห็นว่าเวลาในการประมวลผลแต่ละวิธีมีความแตกต่างกัน โดยที่เวลาในการประมวลผลของวิธี ORELM ใช้การประมวลผลมากที่สุด



รูปที่ 4.9 ระยะเวลาในการคำนวณทางเวลา ELM และ RELM

จากรูปภาพที่ 4.9 แสดงให้เห็นว่าการประมวลผลทางเวลาที่ใช้เทคนิคในการประมาณช่องสัญญาณ โดยวิธี ELM และ RELM มีค่าใกล้เคียงกัน แต่ ค่าอัตราผิดพลาดในการส่งข้อมูลหรือค่าความจุช่องสัญญาณแตกต่างกันส่งผลต่ออนาคตที่จะเลือกใช้งานวิธี RELM เพราะใช้เวลาในการประมวลผลน้อยกว่า ELM โดยที่ RELM นั้นมีประสิทธิภาพดีกว่าเนื่องจากว่าตัวอัลกอริทึมนั้นได้มีการปรับปรุงใหม่เพื่อลดการ Overfitting ของข้อมูลนั่นเอง

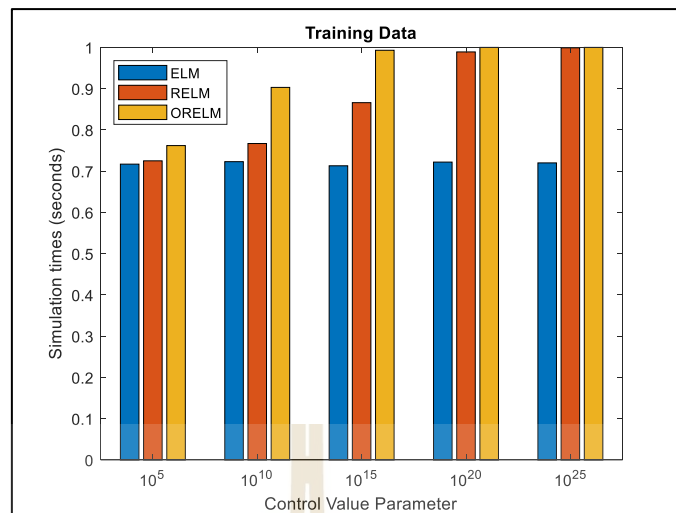


รูปที่ 4.10 ระยะเวลาในการคำนวณทางเวลา ELM, RELM และ ORELM

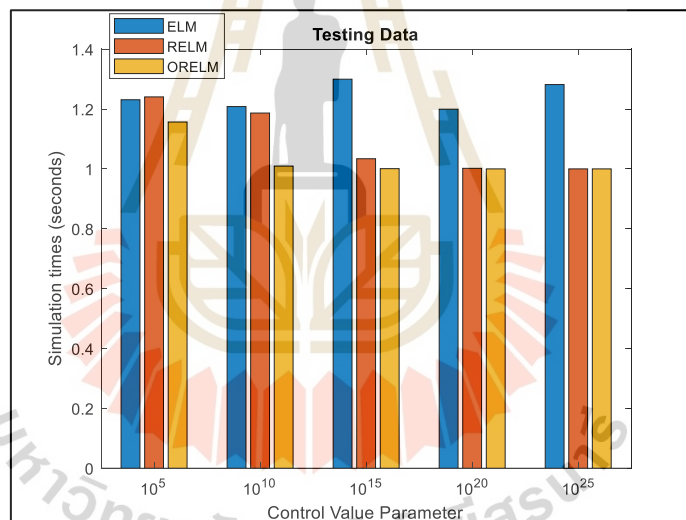
เพื่อเปรียบเทียบให้เห็นความแตกต่างทั้ง 3 วิธีใหม่ที่ผู้วิจัยนำเสนอ รูปภาพที่ 4.10 แสดงให้เห็นว่า การประมวลผลทางเวลาของ ORELM เมื่อเทียบกับ ELM และ RELM ใช้เวลานานที่สุด เพราะด้วยการออกแบบอัลกอริทึมภายในที่มีการเพิ่มตัวแปรในการควบคุมค่าข้อมูลนำเข้า นำออก และการสุ่มค่าต่าง ๆ ภายในโครงสร้าง เมื่อเพิ่มจำนวนโหนด ทีละ 100 ไปจนถึง 500 ซึ่งทำให้เห็นเวลาการคำนวณเพิ่มขึ้น แต่ก็แลกมาด้วยประสิทธิภาพที่สูงขึ้นจากการวัดประสิทธิภาพต่าง ๆ เช่น ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย อัตราความผิดพลาดข้อมูล และความจุช่องสัญญาณ

สำหรับการประเมินประสิทธิภาพการคำนวณทางเวลานั้น หากต้องการแสดงโดยละเอียดทางผู้วิจัยจึงได้ทดสอบเพิ่มเติมโดยการหาผลลัพท์ทางเวลาระหว่าง Training time และตัวแปรในอัลกอริทึมที่มีการควบคุมผลลัพท์โดยแสดงค่าออกมาเป็นเวลา

ซึ่งทางผู้วิจัยจะทำการหาผลลัพท์ทางเวลาโดยวัดค่าจาก 3 อัลกอริทึมคือ ELM, RELM และ ORELM ในช่วงเวลาที่อัลกอริทึมทำการ Training ข้อมูล เพื่อทำการ testing นั้นเอง เพื่อให้สามารถประเมินความเหมาะสมในการนำไปใช้งานหรือปรับปรุงเพิ่มเติมต่อไปดังรูปต่อไปนี้



รูปที่ 4.11 ผลลัพธ์ของการประเมินระหว่างตัวแปรและเวลาในช่วง Training Data



รูปที่ 4.12 ผลลัพธ์ของการประเมินระหว่างตัวแปรและเวลาในช่วง Testing Data

4.6 สรุปผลการดำเนินงาน

จากการเปรียบเทียบผลที่เกิดขึ้น เช่น การประมาณช่องสัญญาณ อัตราความผิดพลาดข้อมูล ความจุช่องสัญญาณ ผลที่เกิดขึ้นแสดงให้เห็นว่า การประมาณช่องสัญญาณ โดยวิธีใหม่ที่ผู้วิจัยได้นำเสนอ ได้แก่ ELM, RELM และ ORELM ให้ประสิทธิภาพที่ดีกว่าระบบเดิม (LS และ MMSE) สามารถนำไปใช้งานระบบโมโฆขนาดใหญ่ในอนาคตได้ และวิธีการ ORELM มีประสิทธิภาพดีที่สุดในแง่ที่จะใช้เวลาในการประมวลผลมากที่สุดเมื่อเทียบกับ ELM และ RELM

บทที่ 5

สรุปเนื้อหาวิทยานิพนธ์และข้อเสนอแนะในการวิจัย

5.1 สรุปเนื้อหาวิทยานิพนธ์

ระบบสื่อสารไร้สายในปัจจุบันมีบทบาทในชีวิตประจำวันเป็นอย่างมากในเรื่องของจำนวนผู้ใช้งานที่มีเพิ่มขึ้นและจากปัญหา เช่น สัญญาณขาดหาย อัตราความเร็วในส่งข้อมูลช้ากว่าความต้องการของผู้ใช้งาน การเพิ่มจำนวนสายอากาศในการสื่อสารและการหาเทคนิคหรือวิธีอื่น ๆ สามารถช่วยพัฒนาระบบหรือแก้ไข้ปัญหาได้ แต่การแก้้ปัญหาอาจแก้ไข้ได้ไม่หมด เพราะในระบบสื่อสารมีจำนวนช่องสัญญาณที่ถูกส่งออกมาจากสถานีฐานมากมาย ดังนั้นงานวิจัยที่จึงนำเสนอโดยมุ่งเน้นศึกษาและพัฒนาประยุกต์ใช้เทคนิคการประมาณช่องสัญญาณกับการเรียนรู้ของเครื่องบนพื้นฐานของระบบการสื่อสารไม่โมฆนานใหญ่ โดยพิจารณาองค์ประกอบที่มีความสำคัญต่องานวิจัยรวมถึงสร้างสถานการณ์จำลองขึ้นมา เพื่อให้ผู้วิจัยได้สามารถประยุกต์ใช้เทคนิคต่าง ๆ เพื่อเปรียบเทียบกับ การประมาณ ช่องสัญญาณ วิธีเดิม เช่น Least Square Estimation (LS) และ Minimum Mean Square Error Estimation (MMSE) โดยเทคนิคที่ผู้วิจัยได้ศึกษาและนำมาประยุกต์ใช้คือกลุ่มของการเรียนรู้ของเครื่อง Extreme Learning Machine (ELM) ซึ่งอัลกอริทึมตัวนี้มี การพัฒนาและปรับปรุงจากหลาย ๆ ศาสตร์และแตกแขนงออกมาเพื่อให้สามารถรองรับกับงานประเภทอื่น ๆ ได้อย่างเหมาะสม อัลกอริทึมมีความเร็วในการเรียนรู้ข้อมูลและทดสอบข้อมูลในระดับวินาที เพราะเป็นโครงข่ายประสาทเทียมแบบ SLFNs จึงง่ายต่อการปรับแต่งพิตเจอร์ต่าง ๆ ซึ่งในการประยุกต์ใช้งานกับระบบสื่อสารนั้น ทางผู้วิจัยได้ทำการศึกษาและเลือกอัลกอริทึมมาทั้งหมด 3 แบบ คือ ELM, RELM และ ORELM การประยุกต์ใช้อัลกอริทึมที่เป็นประเภทการเรียนรู้ของเครื่องกับการประมาณช่องสัญญาณเพื่อทดสอบระบบให้มีประสิทธิภาพการรับส่งข้อมูลดีขึ้น

การประมาณช่องสัญญาณโดยใช้ ELM, RELM และ ORELM ที่เป็นเทคนิคการเรียนรู้ของเครื่องนั้นมีประสิทธิภาพที่ดีกว่าอย่างชัดเจนเมื่อเทียบกับเทคนิคการประมาณช่องสัญญาณแบบ LS และ MSE โดยวิธีการวัดผลในระบบสื่อสาร ได้แก่ ทดสอบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่มีน้อยลง อัตราความผิดพลาดของข้อมูล (BER) ที่น้อยลงอย่างมาก และความจุช่องสัญญาณที่มีเพิ่มขึ้นจากวิธีเดิม โดยอัลกอริทึมที่มีประสิทธิภาพดีที่สุดจากการทดลอง คือ ORELM รองลงมาคือ RELM และ ELM ตามลำดับ โดยที่ ORELM มีการปรับปรุงอัลกอริทึมภายในที่ส่งผลให้การเรียนรู้ข้อมูลและการทดสอบใช้ระยะเวลาที่เพิ่มขึ้น จึงได้ข้อมูลที่มีความเสถียร

มาก ส่วน RELM ก็เป็นอีกอัลกอริทึมที่มีการปรับปรุงเช่นกัน ซึ่งส่งผลให้ข้อมูลลดการ Overfitting ลง และใช้ระยะเวลาที่ในการเรียนรู้และทดสอบน้อยลง โดยสรุปแล้วการประยุกต์ใช้การเรียนรู้ของเครื่องกับการประมวลช่องสัญญาณมีความเป็นไปได้และอาจเป็นอีกทางเลือกในการประมวลช่องสัญญาณในยุคปัจจุบันที่เทคโนโลยี 5G ในอนาคต

5.2 ข้อเสนอแนะและแนวทางการพัฒนาในอนาคต

จากการทดสอบการประมวลช่องสัญญาณกับระบบสื่อสารโมโฆขนานใหญ่ในเชิงทฤษฎีมีความเป็นไปได้ที่จะประยุกต์ใช้งานร่วมกับระบบสื่อสารประเภทอื่น แต่จากการศึกษางานวิจัยพบว่ายังไม่มีระบบสื่อสารหรือเทคนิคต่าง ๆ ที่สามารถประยุกต์ใช้การเรียนรู้ของเครื่องได้นอกเหนือจากอัลกอริทึมเหล่านี้ ดังนั้นการนำเทคนิคการเรียนรู้ของเครื่องไปประยุกต์กับระบบสื่อสารที่ใช้กันในปัจจุบันจึงเป็นอีกแนวทางหนึ่งในการพัฒนางานวิจัยนี้ให้ดียิ่งขึ้นไป



รายการอ้างอิง

- T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wirel. Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010
- C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 303–313, Feb. 2013.
- S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.
- Hien Quoc Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- Gaspar, Guilherme Rodrigues, Montezuma, Paulo. Dinis, Rui, "Channel estimation in massive MIMO systems," *FCT: DEE - Dissertações de Mestrado*, 2016, pp. 28-30.
- J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *1995 IEEE 45th Vehicular Technology Conference. Countdown to the Wireless Twenty-First Century*, 1995, vol. 2, no. 1, pp. 815–819.
- Prince Grover, "5 Regression Loss Functions All Machine Learners Should Know", 2018
- Digital Communications, John Proakis, Massoud Salehi, McGrawHill Education, November 6, 2007
- A. Innok, M. Uthansakul and P. Uthansakul, "The enhancement of MIMO capacity using angle domain processing based on measured channels," *2009 Asia Pacific Microwave Conference*, 2009, pp. 2172-2175, doi: 10.1109/APMC.2009.5385515.
- P. Bartlett, The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network, *IEEE Trans. Inf. Theory* 44 (1998) 525–536.
- Z. Lin, M. Chen, Y. Ma, The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices, *arXiv preprint arXiv:1009.5055* (2010).

- W. Zuo, D. Meng, L. Zhang, X. Feng, D. Zhang, A generalized iterated shrinkage algorithm for non-convex sparse coding, in: The IEEE International Conference on Computer Vision (ICCV), 2013, pp. 217–224.
- Y. Miche, M. Van Heeswijk, P. Bas, O. Simula, A. Lendasse, TROP-ELM: a doubleregularized ELM using LARS and Tikhonov regularization, *Neurocomputing* 74 (2011) 2413–2421.
- J. Yang, Y. Zhang, Alternating direction algorithms for ℓ_1 -problems in compressive sensing, *SIAM J. Sci. Comput.* 33 (2011) 250
- L. Zhang, M. Yang, X. Feng, Y. Ma, D. Zhang, Collaborative representation-based classification for face recognition, *CoRR* (2012), abs/1204.2358.
- K. Amphawan, A. Surarerks and P. Lenca, “Mining periodic-frequent itemset with approximate periodicity using interval transaction-ids list tree”. Proceeding of The 3rd International Conference on Knowledge Discovery and Data Mining: WKDD 2010, Phuket, Thailand, January 9-10.
- Q. Hu, F. Gao, H. Zhang, S. Jin and G. Y. Li, "Deep Learning for Channel Estimation: Interpretation, Performance, and Comparison," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2398-2412, April 2021, doi: 10.1109/TWC.2020.3042074.
- A. L. Ha, T. Van Chien, T. H. Nguyen, W. Choi and V. D. Nguyen, "Deep Learning-Aided 5G Channel Estimation," 2021 15th International Conference on Ubiquitous Information Management and Communication (IMCOM), 2021, pp. 1-7.
- Huang, GB. What are Extreme Learning Machines? Filling the Gap Between Frank Rosenblatt’s Dream and John von Neumann’s Puzzle. *Cogn Comput* 7, 263–278 (2015).



ภาคผนวก ก

บทความที่ได้รับการตีพิมพ์และเผยแพร่

รายชื่อบทความที่ได้รับการตีพิมพ์เผยแพร่ในระหว่างศึกษา

Chittapon Keawin, Peerapong Uthansakul, Apinya Deekaikam. (2021) **ELMx-based Channel Estimation for Massive MIMO Systems** International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology, Pattaya, Thailand, 6-9 May 2009 : pp 362-365.



ELMx-based Channel Estimation for Massive MIMO Systems

Chittapon Keawin
School of Telecommunication Engineering,
Suranaree University of Technology
Muang, Nakhon Ratchasima 30000, Thailand
Email: m6200596@g.sut.ac.th

Peerapong Uthansakul
School of Telecommunication Engineering,
Suranaree University of Technology
Muang, Nakhon Ratchasima 30000, Thailand
Email: uthansakul@sut.ac.th

Apinya Deekaikam
Telecommunications Engineering,
Rajamangala University of Technology Isan
Nakhonratchasima 30000, Thailand
Email: apinya.io@rmuti.ac.th

Abstract—The channel estimation of massive MIMO systems is one of the recent developments in wireless communication technology. The more accurate channel is estimated, the higher performance is achieved. In this paper, the use of machine learning applied to channel estimation is studied. Also, the proposed channel estimations based on Extreme Learning Machine (ELM) family are implemented for massive MIMO systems. These estimations include ELM, Regularized Extreme Learning Machine (RELM) and Outlier Robust Extreme Learning Machine (ORELM). Then, the comparison between two legacy channel estimations using Least Squares (LS) and Minimum Mean Squares Error (MMSE) is presented. The simulation results reveal that the proposed methods significantly overcome LS and MMSE.

Keywords—Channel Estimation, Machine Learning, LS, MMSE, ELM, RELM, ORELM

I. INTRODUCTION

Nowadays, wireless communication is very important. Lots of people are more accessible and competitive with the service providers which has developed 5G [1-2] technology to support the use of many people. It is based on the massive MIMO systems with a greater number of receiving and transmitting antennas. The advantage of massive MIMO is that it is able to extremely deliver fast data. The channel estimation is also an enhancement method for massive MIMO. The widely used channel estimation such as the LS channel estimation and the MMSE channel estimation are both fundamental techniques. However, the LS and MMSE have low accuracy. The LS channel estimation is computationally low of complexity as it does not consider a noise technique. On the other hand, the MMSE take noise in the calculations [3]. In present, solving nonlinear mapping and nonconvex problems has been applied deep learning for channel estimation [4-5]. Also, deep learning has high regression accuracy and faster convergence rate [6]. In the research work presented in [7], the authors have proposed a deep learning for super-resolution channel estimation and DOA estimation based on massive MIMO systems. Although the deep learning technique has better performance, it comes with a longer network training time and complexity in calculating the

channel estimation. Therefore, we propose channel estimation using machine learning three techniques, such as ELM, RELM and ORELM, so called ELMx. The input weight and hidden layer bias are randomly generated from distributions [8-10]. The ELMx is able to gain a faster learning rate and high regression performance. It is used in the training process to optimize the number of hidden neurons. The received signal is taken as input. The simulation results reveal that the channel estimation has an improved performance in terms of Mean Square Error (MSE) and Bit Error Rate (BER).

This paper is organized as follows. In section II, the details of Massive MIMO systems and basic techniques for channel estimation are described. Then, in section III proposes machine learning for channel estimation describing the details of ELMx. In Section IV, we discuss the results. Finally in section V, the conclusion of this paper is given.

II. SYSTEM MODELS

A. Massive MIMO System Model

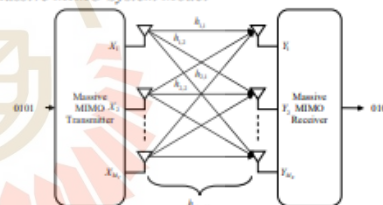


Fig.1 A block diagram of Massive MIMO Systems.

First, we consider a typical massive MIMO system as shown in Fig.1. A block diagram is assumed for sending data from the X (vector of transmitted signals) with M_t transmitting antennas and Y (vector of received signals) with M_r receiving antennas. The relation between transmitted and received signal is given by

$$Y=HX+n \quad (1)$$

Where \mathbf{n} is additive white complex Gaussian noise vector ($M_r \times 1$) and \mathbf{H} is a channel response matrix ($M_r \times M_t$). The relation between transmitted and received signal can be represented in matrix given by

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{M_r} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M_t} \\ h_{21} & h_{22} & \cdots & h_{2M_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_r,1} & h_{M_r,2} & \cdots & h_{M_r,M_t} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{M_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{M_r} \end{bmatrix} \quad (2)$$

In this section, the channel estimation of massive MIMO system is important to transmit multiple sets of data by multiple transmitting antennas which serve to send data in a matrix format, including the various interference signals. For the digital modulation techniques, QPSK, 16QAM, 64QAM, 128QAM and 256QAM are a phase modulation scheme used in constellation mapping. Modulation takes the binary bits as input, is used as a symbol and converts it to complex value. We consider a flat fading MIMO wireless system with M_t transmitting and M_r receiving antennas. The symbol of transmitting by antenna M_t at time instant p is denoted by $X_N(p)$. The transmitted symbols are arranged in the vector of length M_t can be written as

$$X_N(p) = [X_1(p), \dots, X_{M_t}(p)]^T \quad (3)$$

Where $(*)^T$ is the transpose operation of the matrix.

B. LS Channel Estimation

The target of LS channel estimation [11-12] is to minimize the square error distance between the received signal and the estimated one. Therefore, it can find channel estimation \hat{H}_{LS} by

$$\hat{H}_{LS} = \arg \min_{\hat{H}_{LS}} \|Y - \hat{H}_{LS} X\|^2 \quad (4)$$

The channel estimates of impulse responses between all transmitting antennas and receiving antennas are given by

$$\hat{H}_{LS} = YX^H (XX^H)^{-1} \quad (5)$$

where $(*)^H$ is transpose conjugate reserved for the matrix and $(*)^{-1}$ is invert matrix.

C. MMSE Channel Estimation

The MMSE channel estimation [11-14] is the higher accurate version of the LS channel estimation, given by

$$\hat{H}_{MMSE} = \arg \min_{\hat{H}_{MMSE}} \|Y - \hat{H}_{MMSE} X\|^2 \quad (6)$$

For the method of estimating, we consider noise at the expenditure computing time, given by

$$\hat{H}_{(x),MMSE} = YX^H (XX^H + \frac{\sigma_n^2}{\sigma_s^2} I)^{-1} \quad (7)$$

where I is the identity matrix of size $M_r \times M_r$, σ_n^2 is variance noise inversely proportional to the SNR and assuming that every channel response energy is normalized, such as

$$E\{|h_{M_r, M_t}|^2\} = \sigma_h^2 \quad (8)$$

III. MACHINE LEARNING FOR CHANNEL ESTIMATION

Now, we propose an extremely fast learning algorithm for the single hidden layer feedforward networks (SLFNs) with \tilde{N} hidden neurons, where $\tilde{N} \leq N$, the number of training samples.

A. Extreme Learning Machine (ELM)

ELM is one machine learning algorithm that works in the form of a neural networks. It has high regression efficiency and quick learning information, which has been theoretically proven and confirmed. Fig.2 shows the structure of ELM.

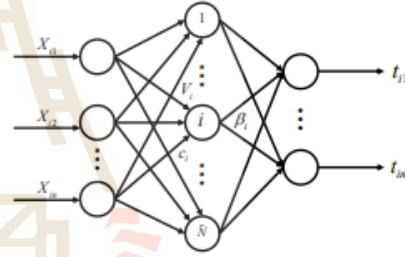


Fig.2 The structure of Extreme Learning Machine.

Considering in the training process of ELM, the investigation is assumed with N training samples (X_i, t_i) , where $X_i = [X_{i1}, X_{i2}, \dots, X_{in}]^T$ and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$. The estimation of SLFNs is modelled by

$$\hat{t}_j = \sum_{i=1}^N \beta_i \phi(c_i X_j + V_i), \quad j = 1, 2, \dots, N \quad (9)$$

where $c_i = [c_{i1}, c_{i2}, \dots, c_{in}]^T$ is input weight vector connecting of the i -th hidden and $\beta = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is output weight vector connecting of the i -th hidden neuron nodes, V_i is the bias

of the i -th hidden neuron nodes, and $o(*)$ is the activation function of the SLFNs. Also, different from the machine learning algorithms, the ELM can randomly generate the input weight c_i and bias V_i .

To improve the performance of regression, one of ability ELM verifies zero error which can approximate all N samples as $\sum_{i=1}^N \|\hat{f}_i - t_i\| = 0, i.e.$

The above N equations can be written compactly as:

$$H\beta = T \quad (10)$$

where

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}_{N \times m} \quad \text{and} \quad T = \begin{bmatrix} T_1 \\ \vdots \\ T_N \end{bmatrix}_{N \times m} \quad (11)$$

$$H(c_1, \dots, c_N, V_1, \dots, V_N, X_1, \dots, X_N) = \begin{bmatrix} o(c_1 \cdot X_1 + V_1) & \dots & o(c_N \cdot X_1 + V_N) \\ \vdots & & \vdots \\ o(c_1 \cdot X_n + V_1) & \dots & o(c_N \cdot X_n + V_N) \end{bmatrix}_{N \times N} \quad (12)$$

where H is the hidden layer output matrix of neural network.

For ELM solution, $\hat{\beta}$, is the smallest norm least-square solution of the linear system given by

$$\hat{\beta} = H^+ T \quad (13)$$

Where $(*)^+$ is the Moore-Penrose pseudoinverse of H .

B. Regularized Extreme Learning Machine (RELM)

The ELM has shown fairly well results in several applications, for a good performance we consider the way of \hat{N} must be made to avoid overfitting and underfitting. In the paper, propose RELM for SLFNs with sigmoid function, is described that smaller norms parameters have can achieve better generalization. The equation can be solving the problem in l_2 norm of β is added, processes to various types of activation functions and hidden neuron nodes, like kernels.

Therefore, RELM can be explained the method is given by

$$\underset{\beta_0, \beta_n, \alpha}{\text{minimize}} \frac{C}{2} \|H\beta + \beta_0 - t\|_2^2 + \frac{(1-\alpha)}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1, \quad (14)$$

where C and β_0 are regularization parameters.

When considering only the l_2 norm penalty ($\alpha = 0$) with $\beta_0 = 0$, the formula of RELM is given by

$$\hat{\beta} = (H^H H + \frac{I}{C})^{-1} H^H t \quad (15)$$

C. Outlier Robust Extreme Learning Machine (ORELM)

Recently, the ORELM is optimized to improve the performance in l_1 norm of methods forms the appearance of outliers.

This can be achieved in ELM used, given by

$$\hat{\beta} = \arg \min_{\beta} \tau \|H\beta - t\|_1, \quad (16)$$

Considering the usual l_2 norm to find the solution of the following optimization is given by

$$\hat{\beta} = \arg \min_{\beta} \tau \|H\beta - t\|_1 + \frac{1}{2C} \|\beta\|_2^2 \quad (17)$$

Finally, after calculating of all channel estimation algorithm, the calculation of Error \hat{H} is estimated that the model predicts the difference from the real H and then find the mean. We find the Gradient of Loss depending on the difference and perform the weight with Backpropagation. Then, we use the Gradient Descent algorithm to reduce losses in the next training round.

The loss function in regression can be called as a mean square error given by

$$MSE_N = \sum_{n=1}^N \|H - \hat{H}\|_2^2 \quad (18)$$

IV. RESULTS AND DISCUSSION

TABLE I. The ELMx channel estimation algorithm

Input:	The scenario simulation massive MIMO system, receive signal vector Y .
Output:	<p>The estimated channel \hat{H} of ELM-ML, RELM-ML, ORELM-ML.</p> <ol style="list-style-type: none"> 1: Firstly, generate simulator to the massive MIMO system and adapt definite noise or distortion into the system. 2: The training sequences are generated, in which receive signal Y are grouped. 3: Optimize the number hidden neuron. 4: Generate real values input weight and bias c_i, V_i randomly. 5: Use the model of hidden layer neuron using (7). 6: Calculate output \hat{H} by (13), (15) and (17). 7: Calculate the MSE performance using (18).

In this section, to validate a group of ELMx algorithms, we need to test the mean square error (MSE), as well as the LS and MMSE channel estimation. The Massive MIMO systems employs 128 transmitting antennas and 128 receiving antennas with QPSK modulation mapping and pilot number. The ELMx channel estimation algorithm is estimated in TABLE I.

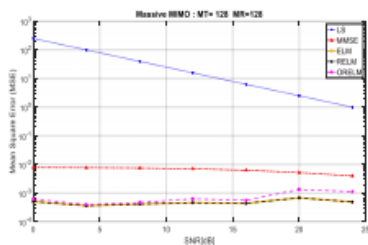


Fig.3 Performance of MSE vs. SNR (dB) for 128x128 MIMO system.

Fig.3, improving the performance of the channel estimation algorithm using MSE concerning various SNR for Massive MIMO is shown. The groups of ELMx are better techniques than LS and MMSE channel estimation. Obviously, MSE decreased as SNR also increased, which was consistent with the target.

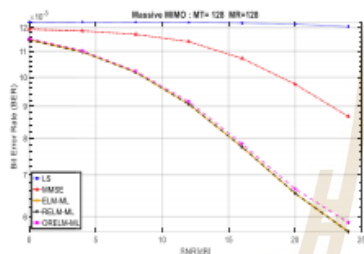


Fig.4 Performance of BER vs. SNR (dB) for 128x128 MIMO system.

Fig.4 shows bit error rate (BER) comparison of LS, MMSE, ELM-ML, RELM-ML and ORELM-ML. Therefore, we consider Massive MIMO based system for communication system. We assume 128 transmitting antennas (M_T) and 128 receive antennas (M_R). As we can see, the groups of ELMx are better techniques than LS and MMSE channel estimation which are basic techniques. At high SNR values, the performance gap is higher than the ones at lower SNR.

V. CONCLUSION

This paper has presented a concept using machine learning in channel estimation employing ELM, RELM and ORELM. They have faster learning and higher regression performance. We have optimized the hidden neurons in the structure for matching the massive MIMO based system. Then, we computed MSE and BER between the basic estimation techniques such as LS and MMSE. From simulation results, the machine learning is better than both techniques. For future work, we have a plan to use auxiliary information-aware ELM where phase [15], linear prediction residual [16], and empirical mode decomposition information [17] may be used as additional information to further improve the performance of conventional

ELM method. Then, we will apply machine learning techniques to 5G communication systems.

ACKNOWLEDGMENT

This work was supported by Suranaree University of Technology (SUT) and Thailand Science Research and Innovation (TSRI).

REFERENCES

- [1] C. I. C. Rowell, S. Han, Z. Xu, G. Li and Z. Pan, "Toward green and soft: a 5G perspective," in *IEEE Communications Magazine*, vol. 52, no. 2, 2014, pp. 66-73.
- [2] Ghavami, Kamran, "Channel Estimation and Symbol Detection In Massive MIMO Systems Using Expectation Propagation," *LSU Doctoral Dissertations*. 4378, 2017.
- [3] Zaib, A., Khattak, S. "Structure-Based Low Complexity MMSE Channel Estimator for OFDM Wireless Systems," *Wireless Pers* 97, 2017, pp. 567-5674.
- [4] Qingchen Zhang, Laurence T. Yang, Zhikui Chen, Peng Li, "A survey on deep learning for big data" *Information Fusion*, 2018, pp. 146-157.
- [5] LeCun, Y., Bengio, Y. & Hinton, G. "Deep learning," *Nature* 521, 2015, pp.436-444.
- [6] C. Luo, J. Ji, Q. Wang, X. Chen and P. Li, "Channel State Information Prediction for 5G Wireless Communications: A Deep Learning Approach," in *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, 2020, pp. 227-236.
- [7] H. Huang, J. Yang, H. Huang, Y. Song and G. Gui, "Deep Learning for Super-Resolution Channel Estimation and DOA Estimation Based Massive MIMO System," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, 2018, pp. 8549-8560.
- [8] W. Deng, Q. Zheng and L. Chen, "Regularized Extreme Learning Machine," 2009 *IEEE Symposium on Computational Intelligence and Data Mining*, Nashville, TN, 2009, pp. 389-395.
- [9] G. Huang, H. Zhou, X. Ding and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," in *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, 2012, pp. 513-529.
- [10] Zhang, K., & Luo, M. "Outlier-robust extreme learning machine for regression problems. *Neurocomputing*," vol. 151, 2015, pp. 1519-1527.
- [11] Gaspar, Guilherme Rodrigues, Montezuma, Paulo, Dinis, Rui, "Channel estimation in massive MIMO systems," *FCT: DEE - Dissertações de Mestrado*, 2016, pp. 28-30.
- [12] Promsuvana N., Uthansakul P., "Feasibility of adaptive 4x4 MIMO system using channel reciprocity in FDD mode," 14th *Asia-Pacific Conference on Communications, APCC 2008*, pp. 1-6.
- [13] Innok A., Uthansakul P., Uthansakul M., "Angular beamforming technique for MIMO beamforming system," *International Journal of Antennas and Propagation*, 2012, pp. 1-9.
- [14] Uthansakul P., Bialkowski M.E., "Multipath signal effect on the capacity of MIMO MIMO-OFDM and spread MIMO-OFDM," 15th *International Conference on Microwaves, Radar and Wireless Communications, MIKON - 2004*, pp. 989-992.
- [15] K. Phapatanaburi, P. Buayai, W. Naktong, and J. Srinonchat, "Exploiting Magnitude and Phase Aware Deep Neural Network for Replay Attack Detection," *ECTI-EEC*, vol. 18, no. 2, 2020, pp. 89-97.
- [16] K. Phapatanaburi, P. Buayai, M. Kupimai and T. Yodrot, "Linear Prediction Residual-Based Constant-Q Cepstral Coefficients for Replay Attack Detection," 2020 *8th International Electrical Engineering Congress (iEECON)*, Chiang Mai, Thailand, 2020, pp. 1-4.
- [17] K. Phapatanaburi, K. Kokkhunthod, L. Wang, T. Jumphoo, M. Uthansakul et al., "Brainwave classification for character-writing application using emd-based gmm and kelm approaches," *Computers, Materials & Continua*, vol. 66, no.3, 2021, pp. 3029-3044.

ประวัติผู้เขียน

นาย ชิตกณ แก้วอินทร์ เกิดเมื่อวันที่ 1 ธันวาคม พ.ศ. 2539 เริ่มศึกษาชั้นอนุบาลที่โรงเรียนอนุบาลนครราชสีมา ชั้นอนุบาลศึกษาปีที่ 1-2 ที่โรงเรียนอนุบาลนครราชสีมา ชั้นประถมศึกษาปีที่ 1-6 ชั้นมัธยมศึกษาปีที่ 1-3 ที่โรงเรียนราชสีมาวิทยาลัย ชั้นระดับประกาศนียบัตรวิชาชีพ (ปวช.) ศึกษาปีที่ 1-3 วิทยาลัยเทคนิคนครราชสีมา จังหวัดนครราชสีมา และสำเร็จการศึกษาระดับปริญญาตรี สาขาวิชาวิศวกรรมโทรคมนาคม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี จังหวัดนครราชสีมา เมื่อปี พ.ศ.2561

ปี พ.ศ.2562 เข้าศึกษาต่อในระดับปริญญาโท สาขาวิชาวิศวกรรมโทรคมนาคม มหาวิทยาลัยเทคโนโลยีสุรนารี โดยขณะศึกษาได้รับทุนการศึกษาแก่นักศึกษาระดับบัณฑิตศึกษาที่ผ่านการคัดเลือกให้เป็นผู้ได้รับทุนวิจัย ประเภททุนการศึกษาสำหรับผู้มีศักยภาพเข้าศึกษาระดับบัณฑิตศึกษา ประจำปีการศึกษา 2562

ผลงานวิจัย : ได้เสนอบทความเข้าร่วมในการประชุมวิชาการระดับนานาชาติ 2021 IEEE 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON 2021), Chiang Mai, Thailand วันที่ 19 พฤษภาคม 2564 เรื่อง ELMx-based Channel Estimation for Massive MIMO Systems.

มหาวิทยาลัยเทคโนโลยีสุรนารี