

Performance Estimation

สมาชิกกลุ่ม

นายกฤตพล ผ้าเจริญ 6401012620161

นายเจษฎา ศรีจุลโพธิ์ 6401012620170

นายศุภกร พลศิริ 6401012620234

นายสิรภพ ห่วงวิไล 6401012630132

ข้อที่ 1

ใช้ data set ที่เรียน ทุกการทดลองให้เพิ่มจำนวนครั้งในการทดลองซ้ำให้มากกว่าใน lecture และรายละเอียดอื่นๆเพิ่มเติมที่คิดว่าจำเป็น

การทดสอบความเที่ยงตรง (Precision) วิธี Holdout

Holdout จำนวน 20 คน

20 seeds

-----10%-----
Average RMSE : 27.5366
Standard Deviation: 57.4182

-----20%-----
Average RMSE : 9.2948
Standard Deviation: 2.5371

-----30%-----
Average RMSE : 8.4142
Standard Deviation: 1.8084

-----40%-----
Average RMSE : 8.1178
Standard Deviation: 1.3866

-----50%-----
Average RMSE : 7.8889
Standard Deviation: 1.2278

-----60%-----
Average RMSE : 7.7964
Standard Deviation: 1.1417

-----70%-----
Average RMSE : 7.5731
Standard Deviation: 1.3443

-----80%-----
Average RMSE : 7.4505
Standard Deviation: 1.4258

-----90%-----
Average RMSE : 7.4638
Standard Deviation: 2.1044

การทดสอบความเที่ยงตรง (Precision) ของวิธี k- fold Cross-Validation

k-fold Cross-Validation จำนวน 20 คน

20 seeds

=====2-fold=====

Average RMSE: 7.8784

Standard Deviation: 0.9597

=====5-fold=====

Average RMSE: 7.3804

Standard Deviation: 0.3220

=====3-fold=====

Average RMSE: 7.5800

Standard Deviation: 0.6298

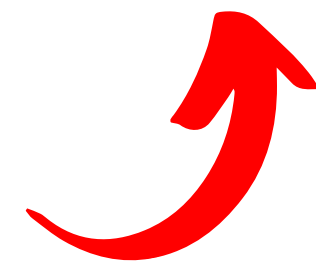
=====10-fold=====

Average RMSE: 7.1651

Standard Deviation: 0.1942

เปรียบเทียบ Holdout และ Cross-Validation

	RMSE	SD
Holdout 60%	7.7964	1.1417
10-fold CV	7.1651	0.1942



**การทดสอบความเที่ยงตรง (Precision) และความแม่นยำ
(Accuracy) เมื่อจำนวนข้อมูลมากขึ้น**

Holdout จำนวน 100 คน

100 seeds

-----10%-----

Average RMSE : 6.3293
Standard Deviation: 0.5321

-----20%-----

Average RMSE : 6.0292
Standard Deviation: 0.3498

-----30%-----

Average RMSE : 5.9283
Standard Deviation: 0.3419

-----40%-----

Average RMSE : 5.8755
Standard Deviation: 0.3515

-----50%-----

Average RMSE : 5.8319
Standard Deviation: 0.4027

-----60%-----

Average RMSE : 5.8103
Standard Deviation: 0.4989

-----70%-----

Average RMSE : 5.8208
Standard Deviation: 0.6100

-----80%-----

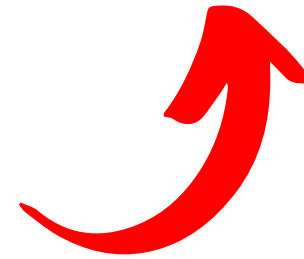
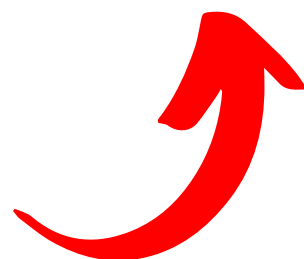
Average RMSE : 5.8753
Standard Deviation: 0.7545

-----90%-----

Average RMSE : 5.7289
Standard Deviation: 1.2668

เปรียบเทียบ Holdout 20 และ 100 คน

Holdout	20 คน		100 คน	
	RMSE	SD	RMSE	SD
20%	9.2948	2.5371	6.0292	0.3498
50%	7.8889	1.2278	5.8319	0.4027
80%	7.4505	1.4258	5.8753	0.7545



k-fold Cross-Validation จำนวน 100 คน

100 seeds

=====2-fold=====

Average RMSE: 5.8789

Standard Daviation: 0.1509

=====3-fold=====

Average RMSE: 5.8357

Standard Daviation: 0.1021

=====5-fold=====

Average RMSE: 5.7974

Standard Daviation: 0.0687

=====10-fold=====

Average RMSE: 5.7183

Standard Daviation: 0.0626

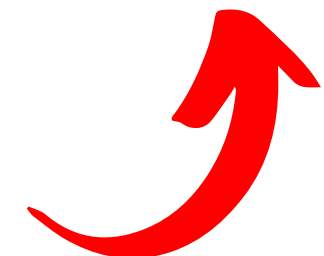
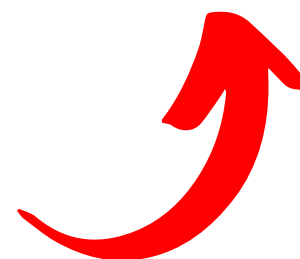
=====20-fold=====

Average RMSE: 5.5749

Standard Daviation: 0.0704

เปรียบเทียบ Cross-Validation จำนวน 20 และ 100 คน

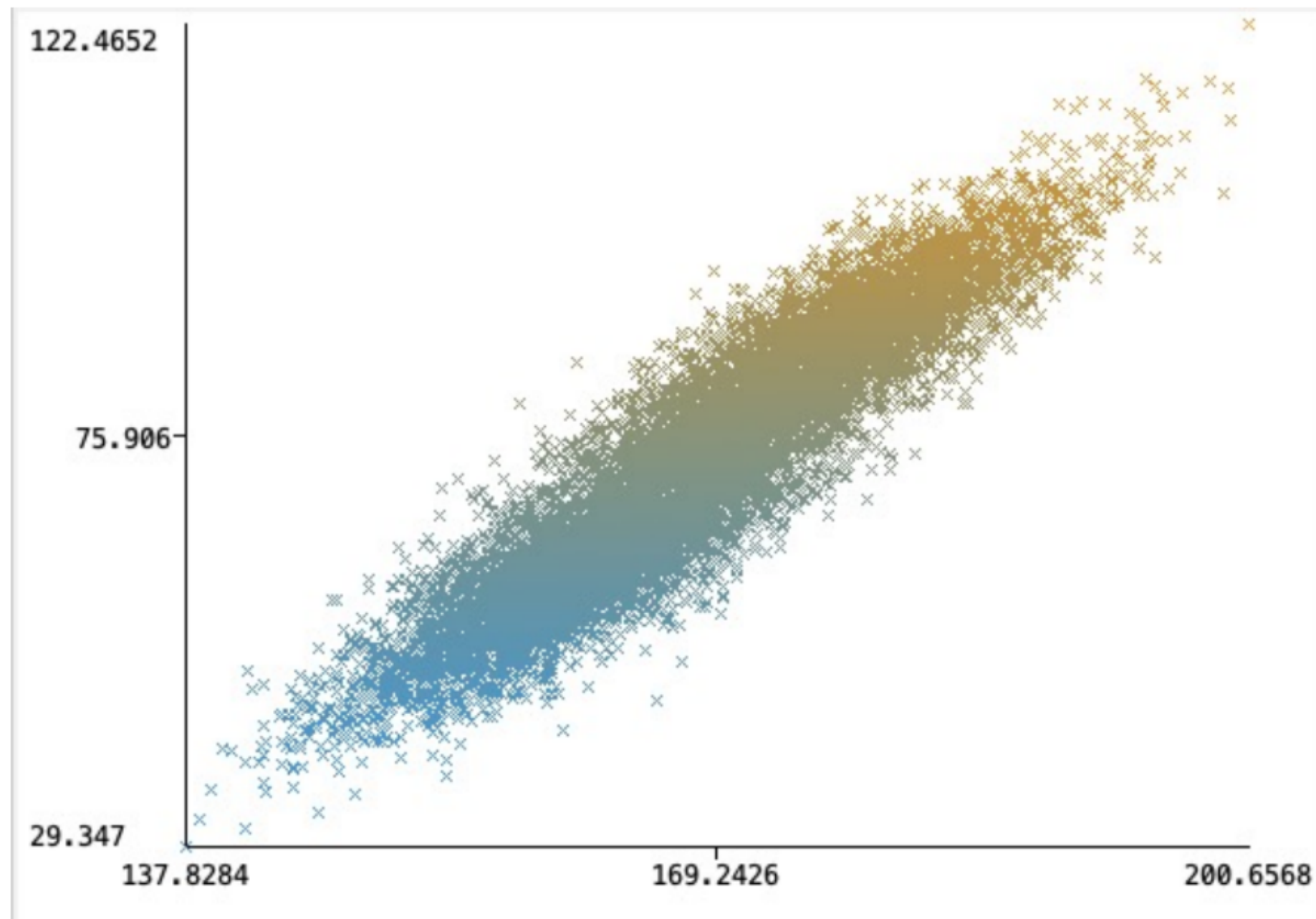
Cross-Validation	20 คน		100 คน	
	RMSE	SD	RMSE	SD
2 fold	7.8784	0.9597	5.8789	0.1509
5 fold	7.3804	0.3220	5.7974	0.0687
10 fold	7.1615	0.1942	5.7183	0.0625



การทดสอบความแม่นยำ (Accuracy) ของเครื่องมือ วัด ประสิทธิภาพของแบบจำลอง

Reference performance

ข้อมูลส่วนสูงและน้ำหนักของ 10,000 คน



Weight = $1.3781 \times \text{Height} + -159.0916$
Average RMSE: 5.5422

Reference RMSE 5.5422

	Training	10-fold CV	Holdout 50%
RMSE	5.2317	5.2218	6.0118
SD	0.8094	0.8225	1.0651

สุ่มข้อมูลออกมา 20 ตัวอย่างจากข้อมูล 10,000 ตัวอย่างด้วย Seed ที่แตกต่างกัน (100 seeds)

การทดลองเพิ่มเติม

**ใช้ข้อมูลทั้งหมด 10,000 ตัวอย่างสร้างแบบจำลอง
Linear Regression โดยเลือก Test options เป็น Use
training set, Holdout และ Cross-validation สังเกต
ค่า RMSE และอภิปรายผล**

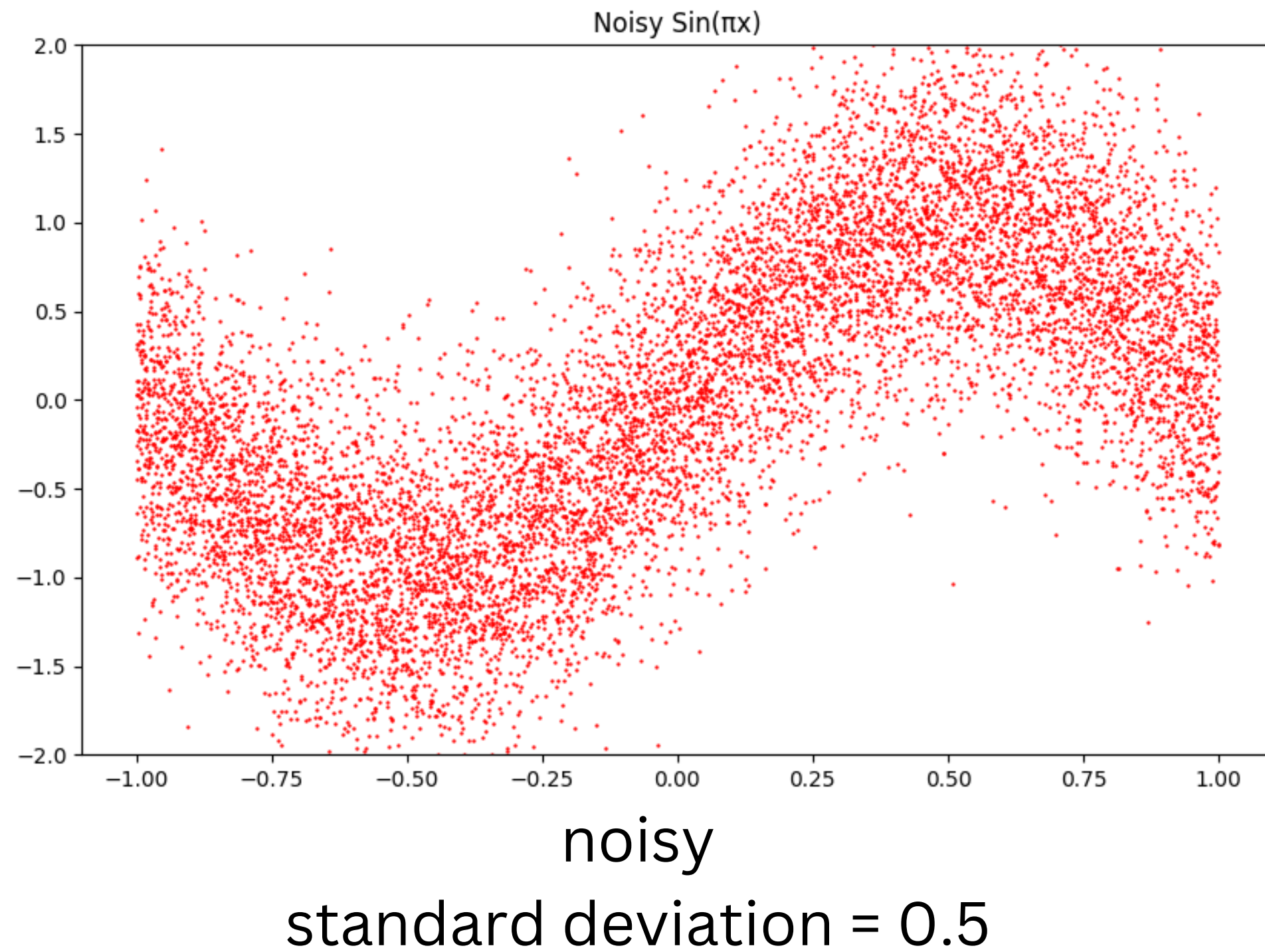
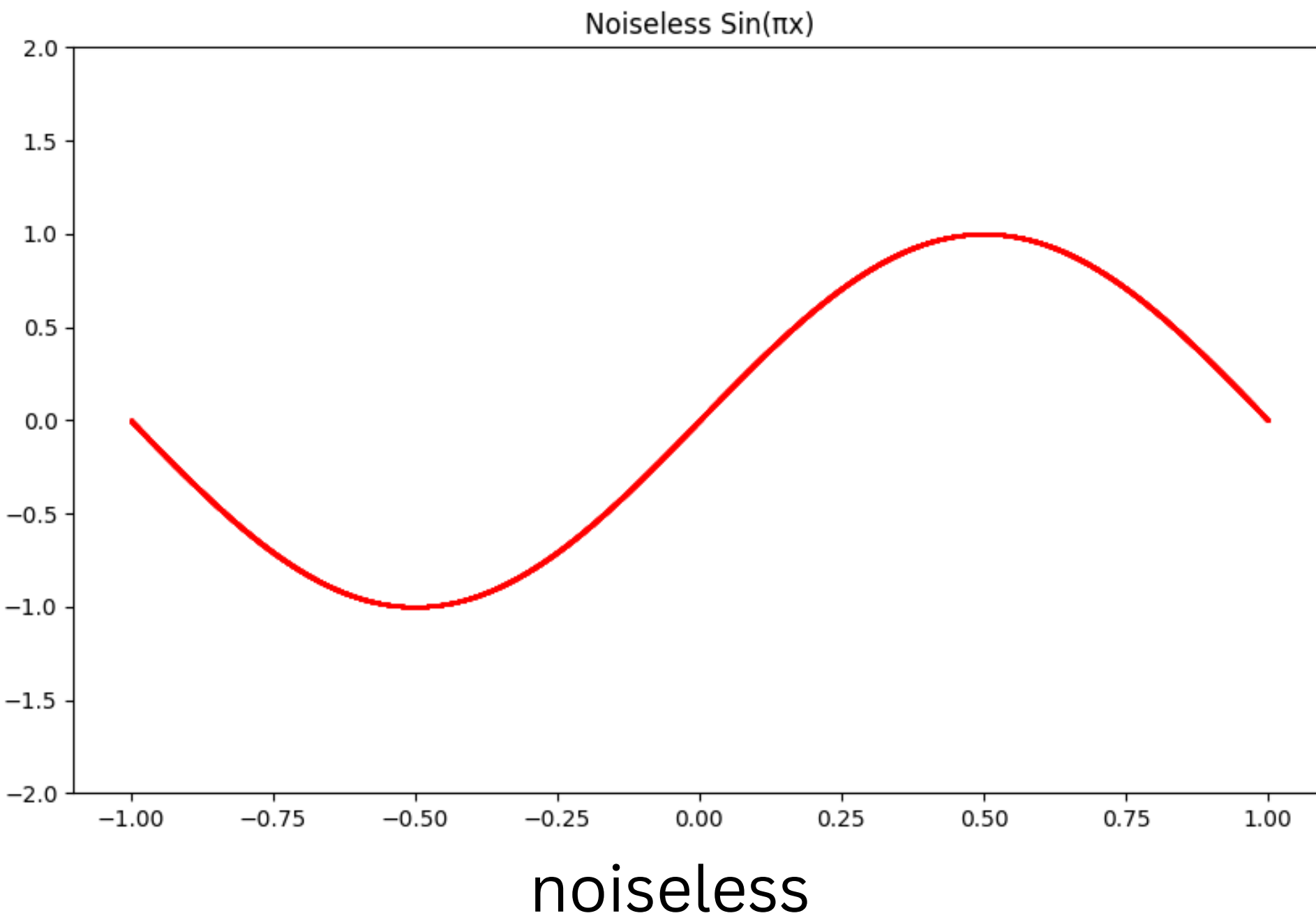
ผลลัพธ์ที่ได้

	Training	10-fold CV	Holdout 50%
RMSE	5.5422	5.542	5.5458
SD	0	0.0006	0.0459

ข้อที่ 2

**Generate Data ขึ้นมาเอง เช่น noiseless หรือ noisy
ของฟังก์ชัน \sin เพื่อสร้างกราฟ Expected Validation
Error ตาม lecture หน้า 60**

ข้อมูล \sin ทั้งหมด 10,000 ค่า



n ครั้ง

สุ่มข้อมูล 20 ข้อมูลจากข้อมูล
ทั้งหมด

16 ครั้ง

กำหนดจำนวนของข้อมูล test
(เปลี่ยนค่า K)

m ครั้ง

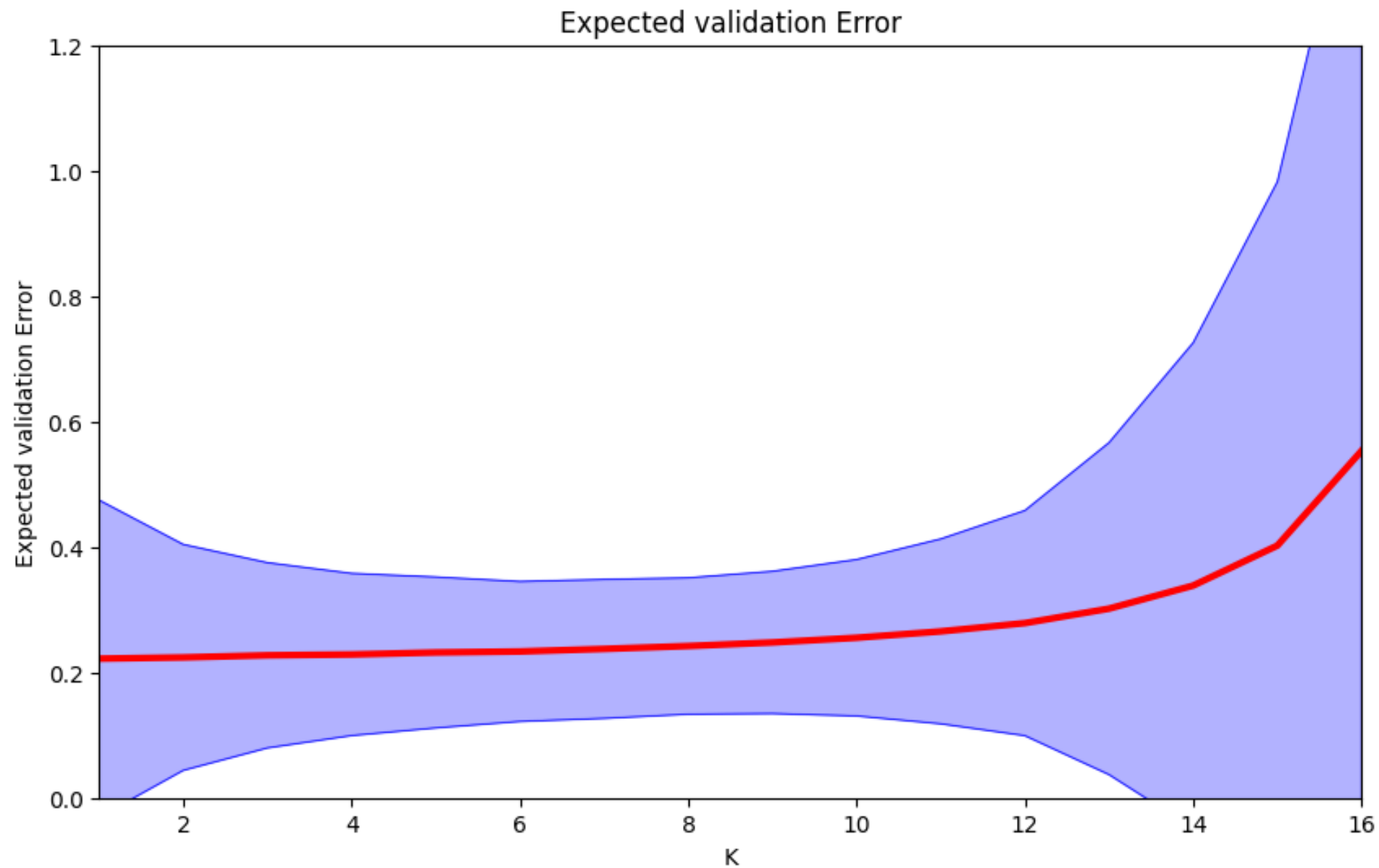
สลับข้อมูล
train test

เก็บข้อมูล error

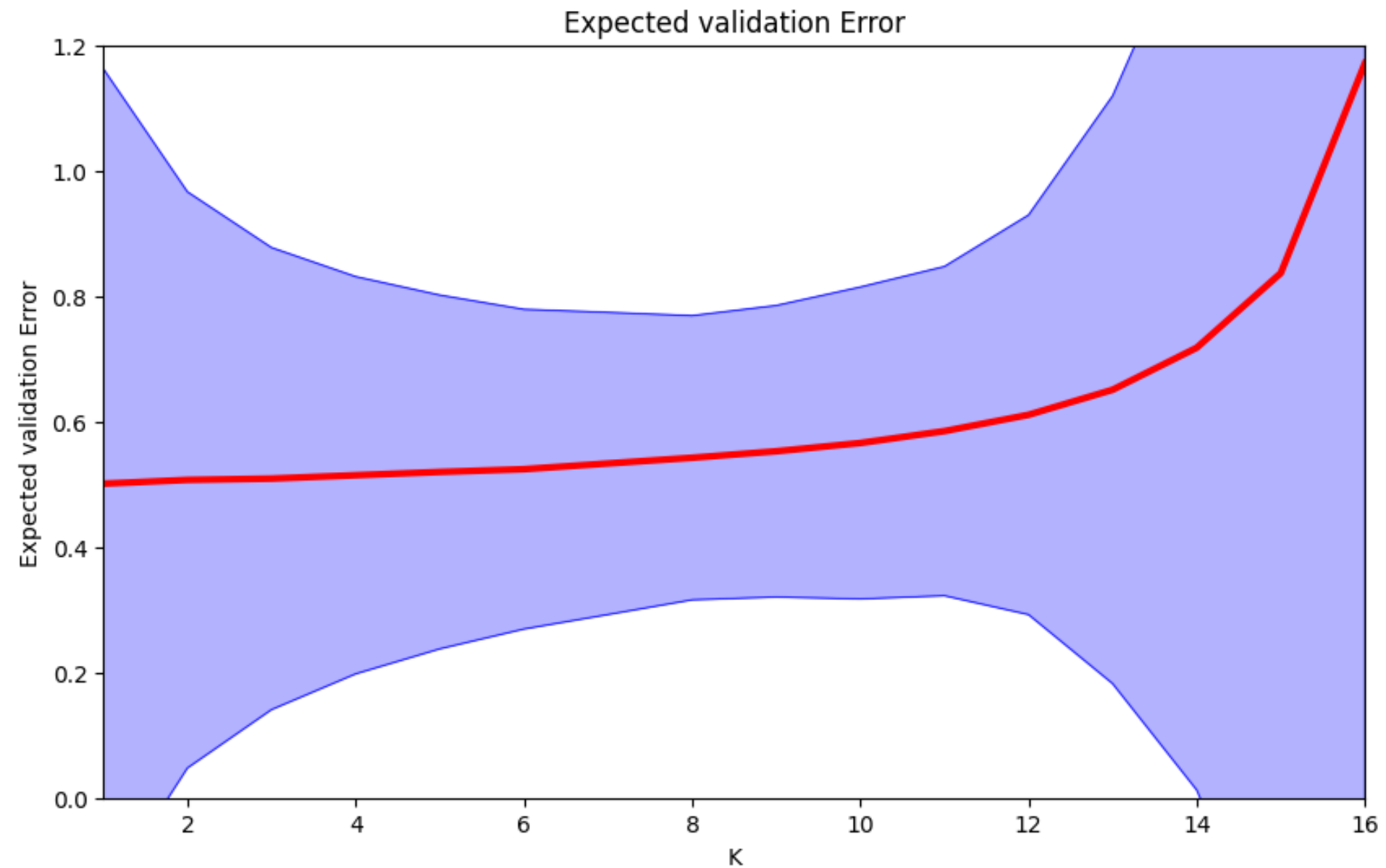
End Loop

N = 20

100 seeds, 500 sample



noiseless



noisy

standard deviation = 0.5