BIRD MIGRATION

Neurale Netze Exam project Presentation 2024/2025

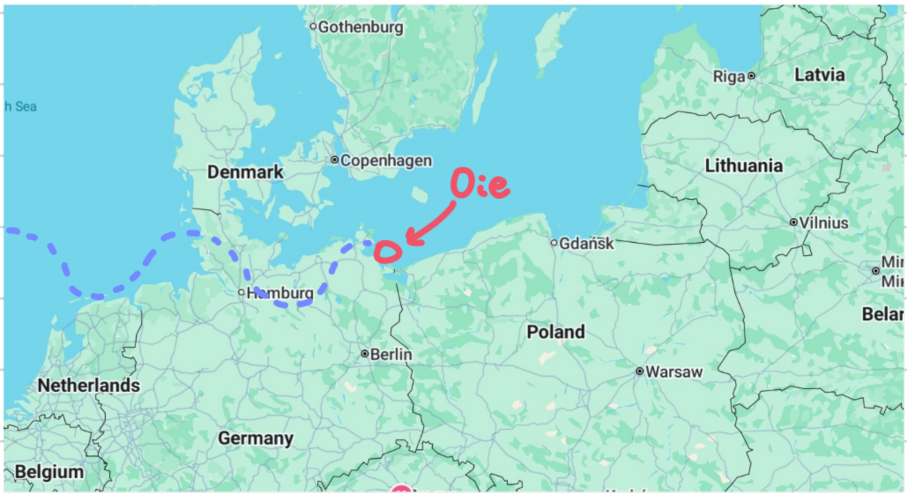
Made by Lorenzo Borghi

INDEX:

CONTEXT

GreifsWalder Oie is an island in the Baltic Sea.

Birds migrate here!



Our goal is to predict number of birds arriving, one day ahead, and distinguishing between spring and autumn.

DATA ANALYSIS

Number of input Features: **101** columns!!

- CALENDAR
- WEATHER

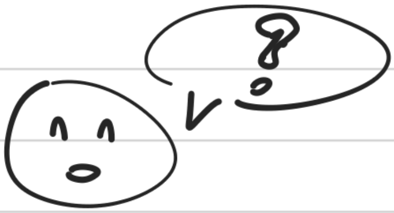
TIMESTAMP	YEAR	MONTH_COS	MONTH_SIN	DAY_COS	DAY_SIN	DOY_SIN	DOY_SIN

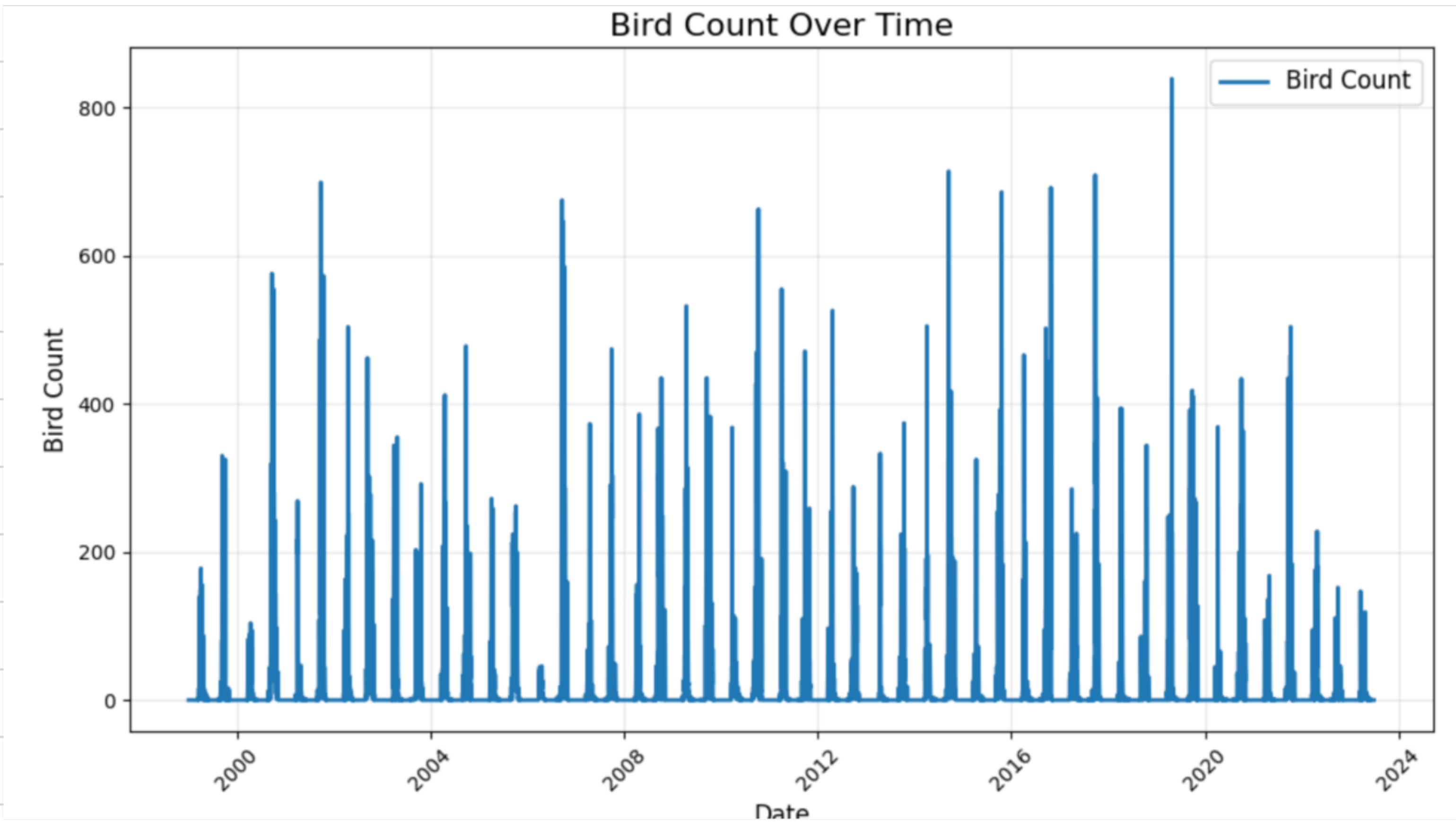
TEMP_Oie	P_Oie	PPN_Oie	WS_Oie	TEMP_DE1	P_DE1	PPN_DE1	...

...	CC_DE1	WS_DE1	WS_DE1	TEMP_DE2	P_DE1	PPN_DE2	...

TARGET = $\boxed{\text{COUNT}}$

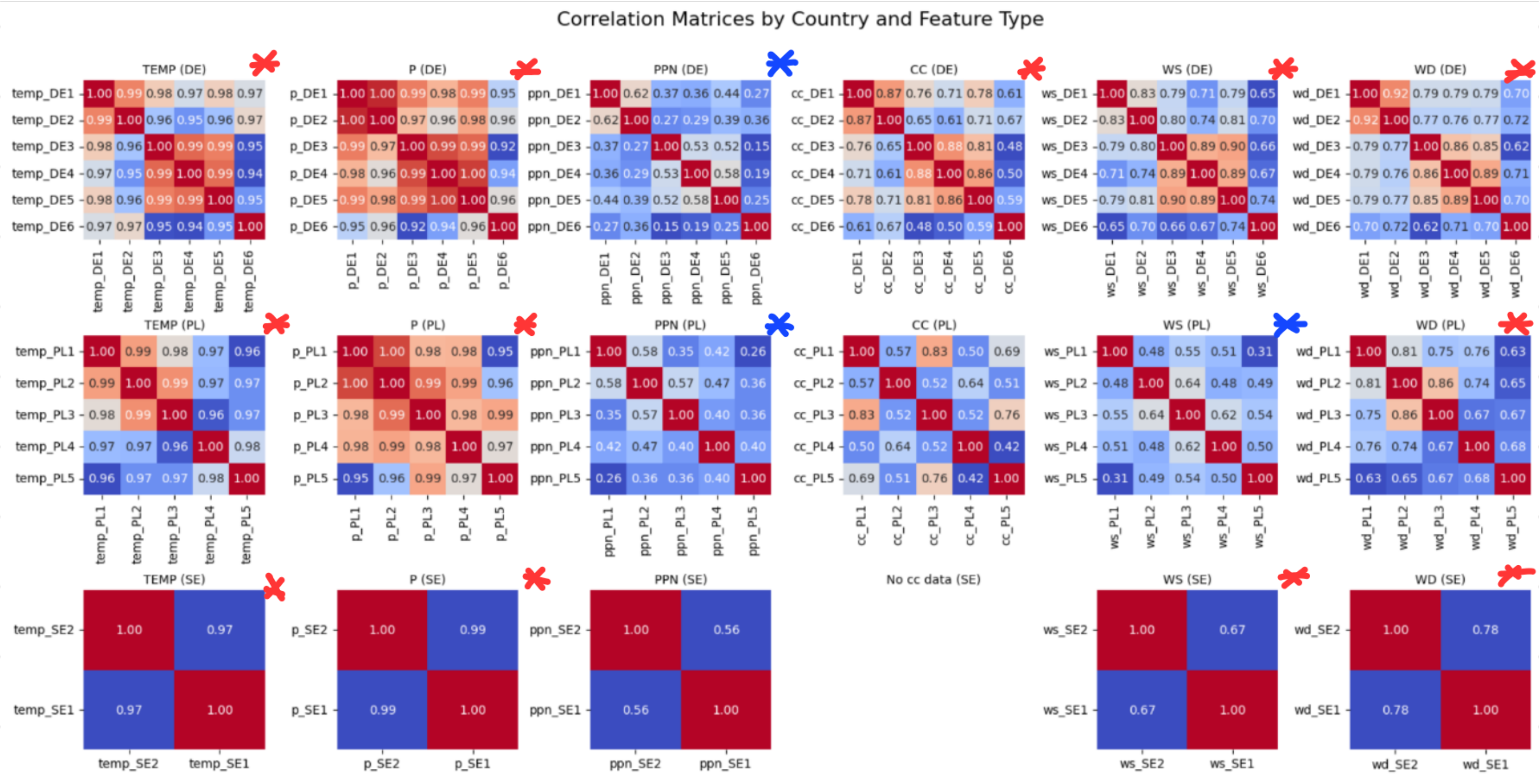
BIRDS BEHAVIOUR





Can we cut some columns? => Less Redundancy

1. Analyze corelation between national data

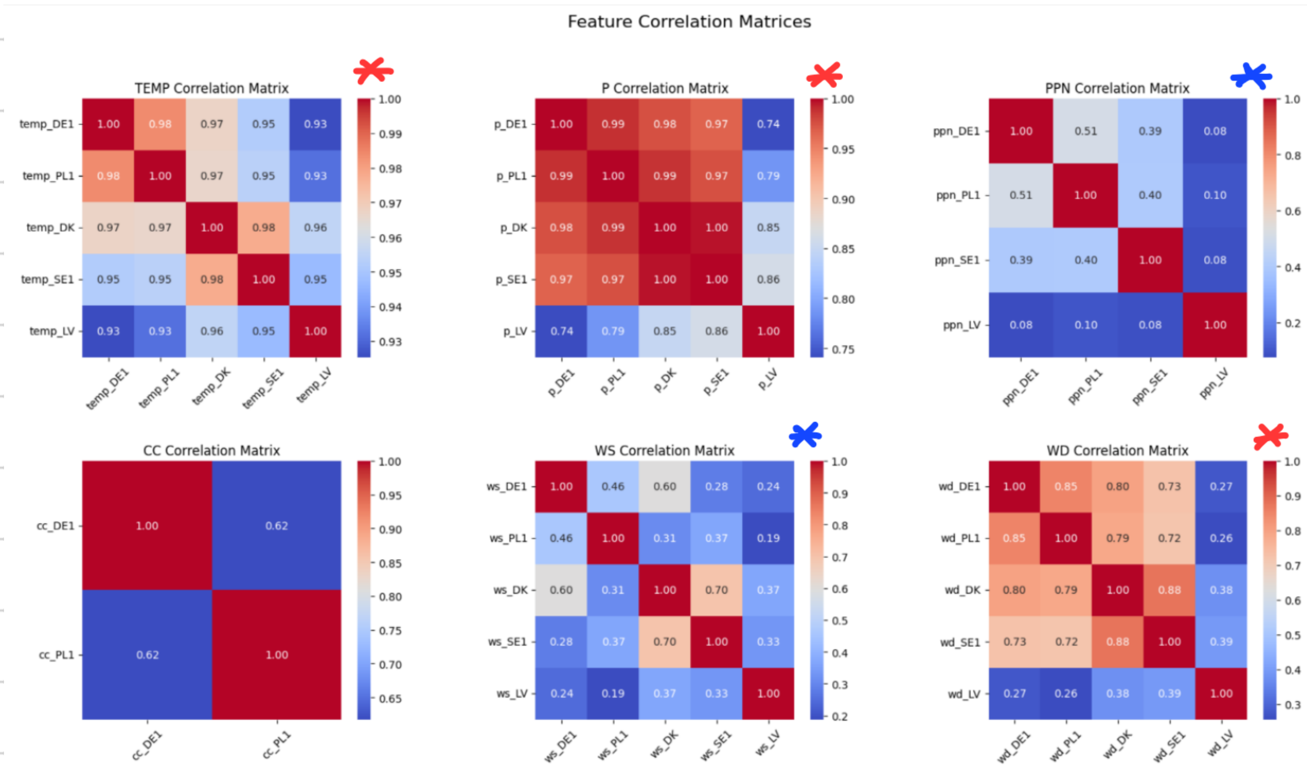


My considerations: • High corelation > 0.7 (*)

• Low corelation < 0.5 (*)

Overall situation shows high redoundancy of international data!!

2. Redoundancy over international data (across countries)

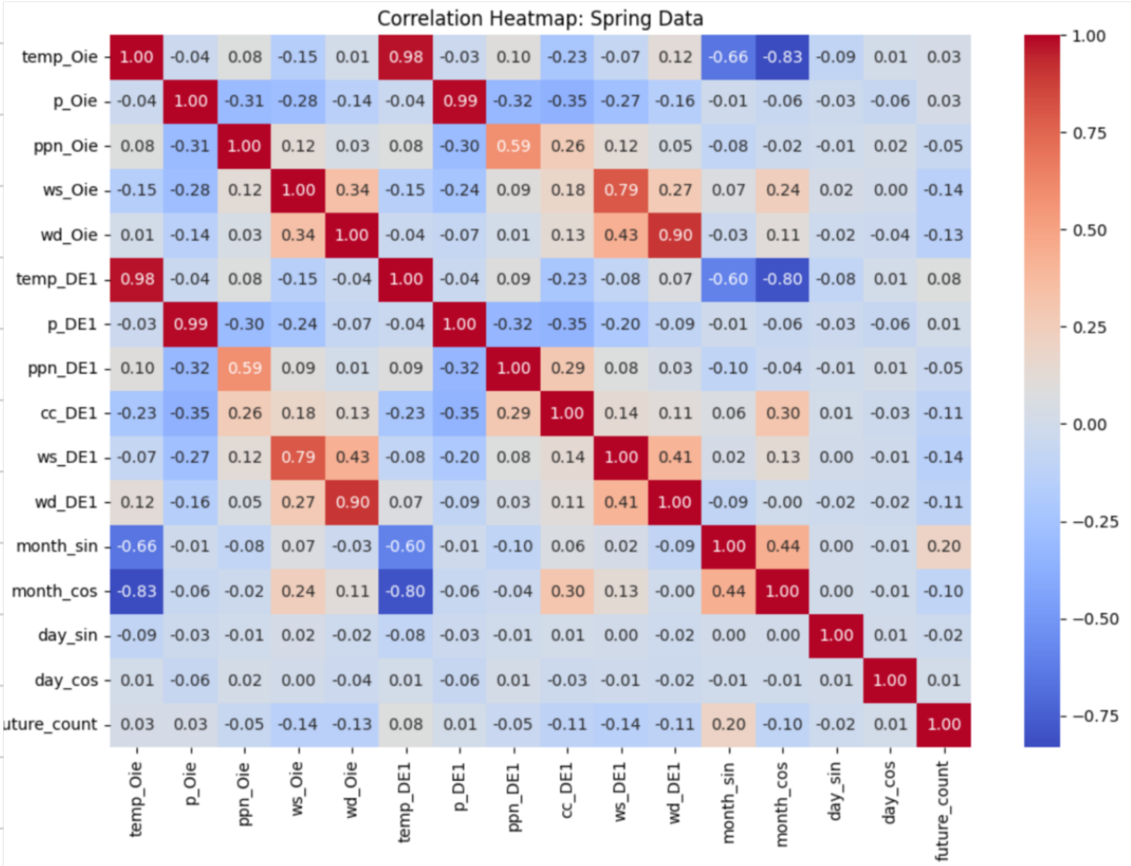


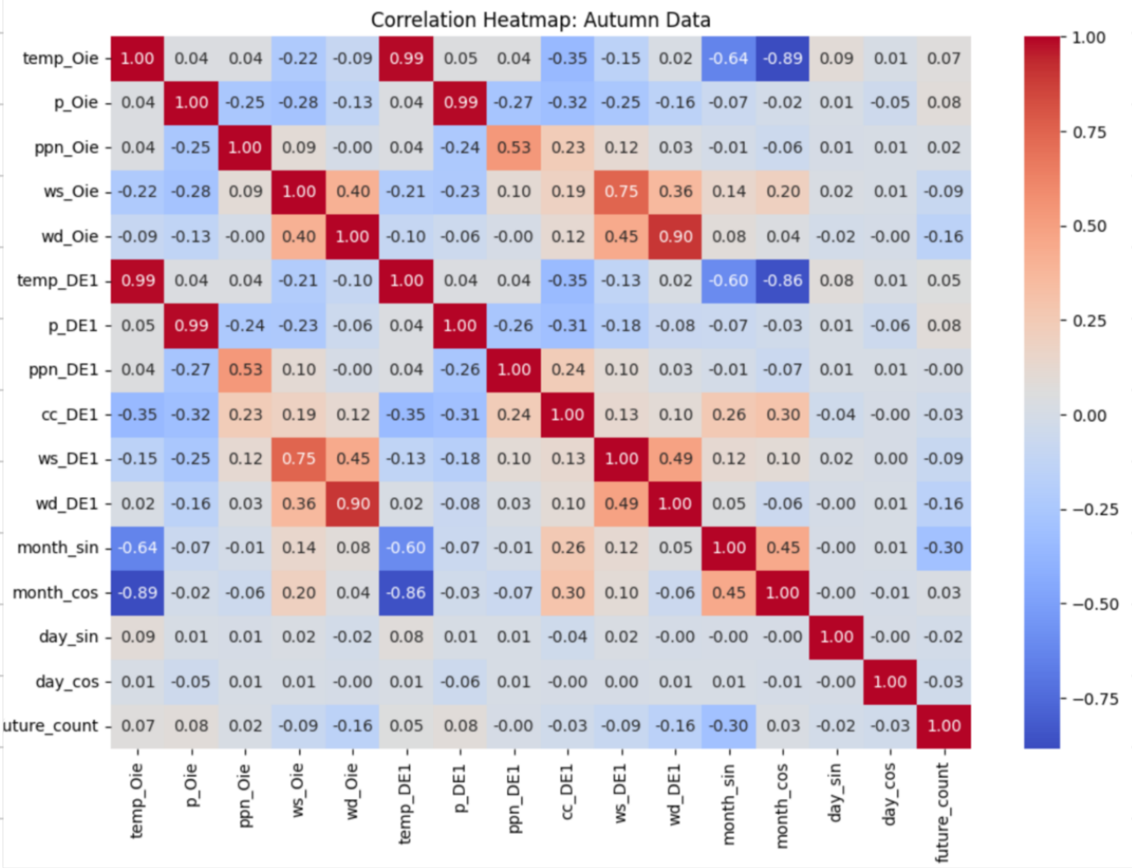
Here we have a similar situation.

3. About Calendar Features, I'm using: | MONTH_COS | MONTH_SIN | DAY_COS | DAY_SIN |

Why?
- YEAR => represents long term periods.
- SPRING => used to separate data sets.

4. We finally analyze which features may have more important results:





Spring: month, temp_0:e and day
Autumn: p_0:e, temp_0:e and month

Anyway NN may reveal us new Non-linear relations...

Predictors:

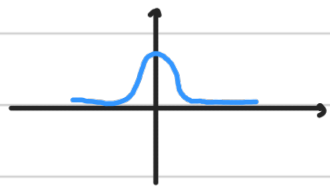
TEMP_0:e	P_0:e	PPN_0:e	WS_0:e	WD_0:e	PPN_DE1	CC_DE1

MONTH_SIN	MONTH_COS	DAY_SIN	DAY_COS

DATA MODELING:

80% TRAIN	20% TEST

Normal distribution



Are Training and Test set homogeneus? =>

	AUT	SPR
MEAN TRAIN	22.35	29.47
MEAN TEST	24.20	73.57
STD TRAIN	12.12	44.95
STD TEST	12.67	47.50

OK!

BENCHMARK MODEL

We generate a Linear Regression both For Spring and Autumn.

The two criteria of evaluation of models are: RMSE and R^2 SCORE

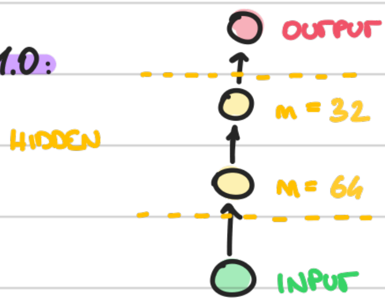
Linear Model

	SPRING	AUTUMN
RMSE	44.90	68.71
R^2	0.11	0.13

NEURAL NETWORK

Creating a basic NN model to quicly test NN task capabilities.

Model 1.0:

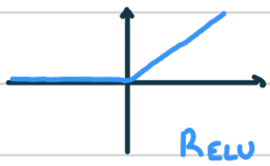


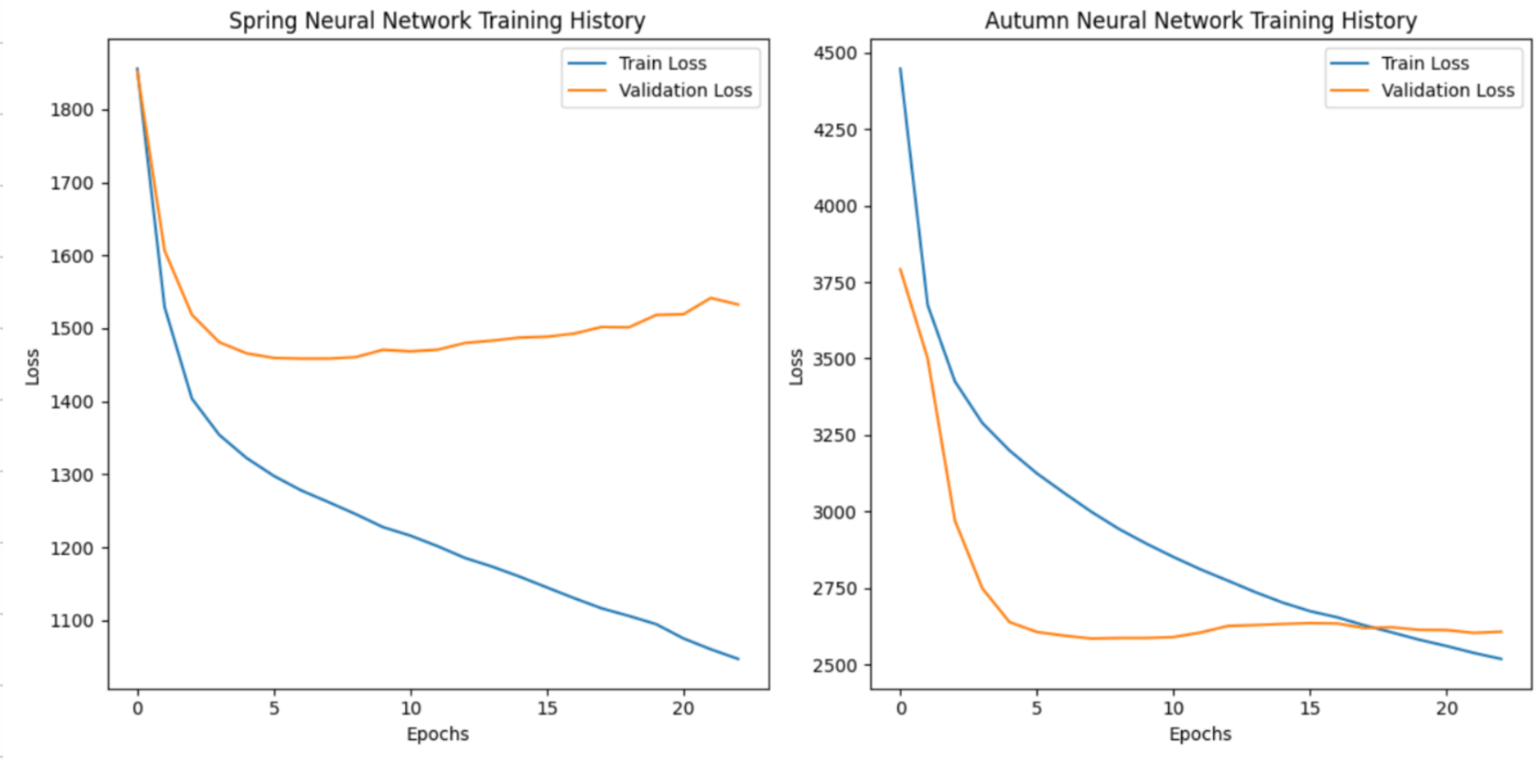
ACTIVATION FUN: RELU ; LOSS: MSE

LEARNING RATE: ADAM, $\eta = 0.005$

EPOCHS: 100

BATCH SIZE: 30

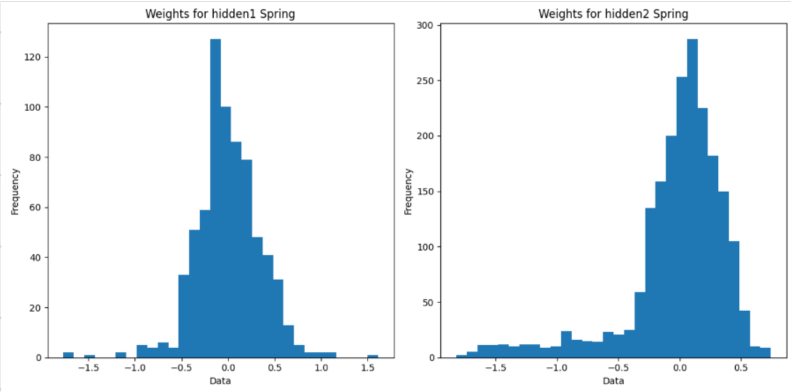


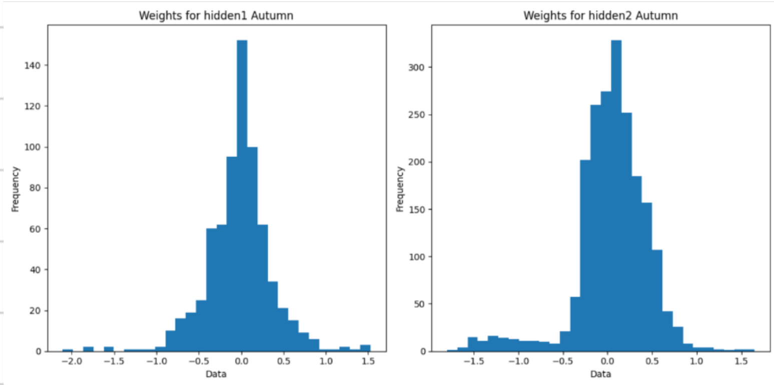


	SPRING	AUTUMN
RMSE	38.19⁺	59.25⁺
R^2	0.35⁺	0.35⁺

Impressive results as a first model, compared to Benchmark!
Reveals that relations involved on "count" predictions are Highly NON-LINEAR

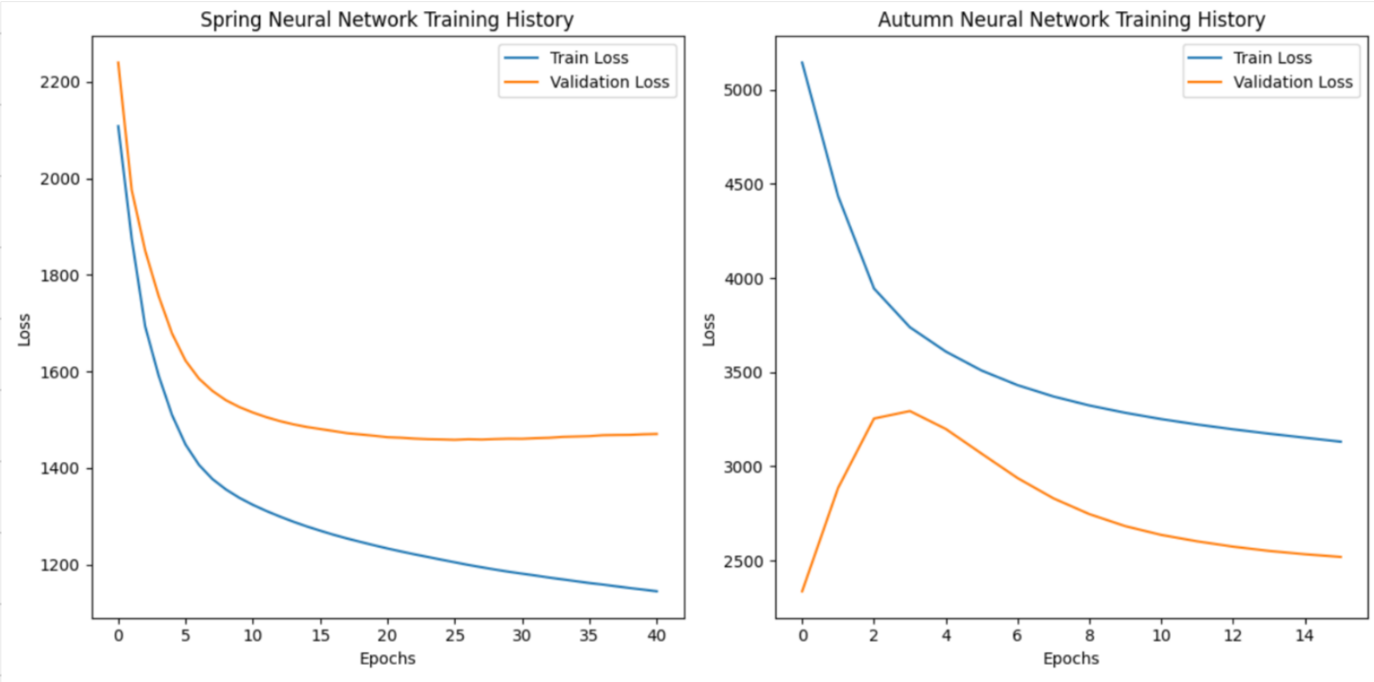
The graphic shows overfitting. Let's inspect the weights:



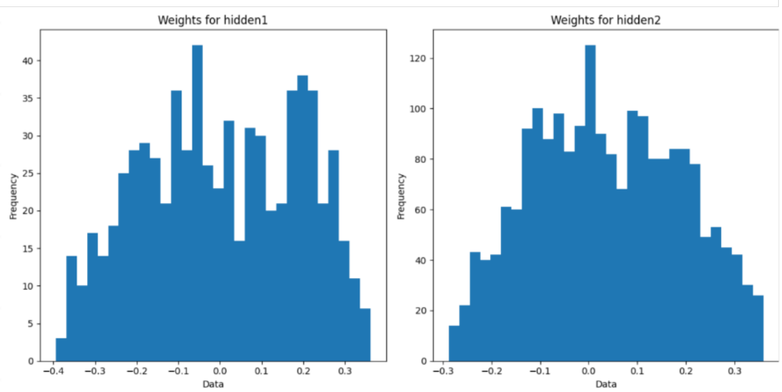


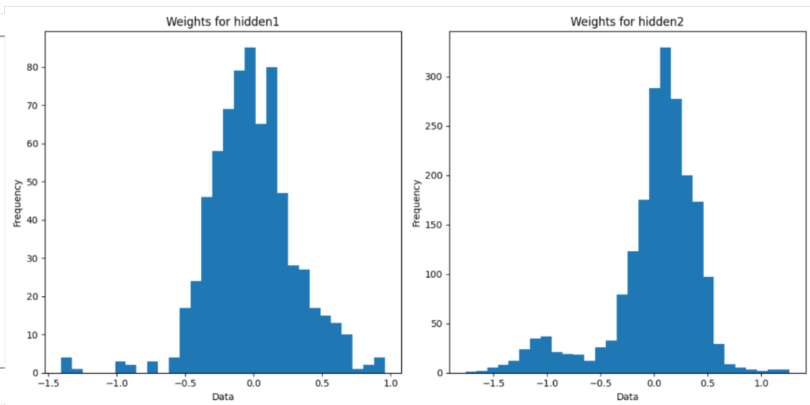
I now apply square weight decay on both hidden layers an reduce LR:

L1L2 REGULARIZERS: $\mu = 0.001$; LEARNING RATE: $\mu = 0.001$

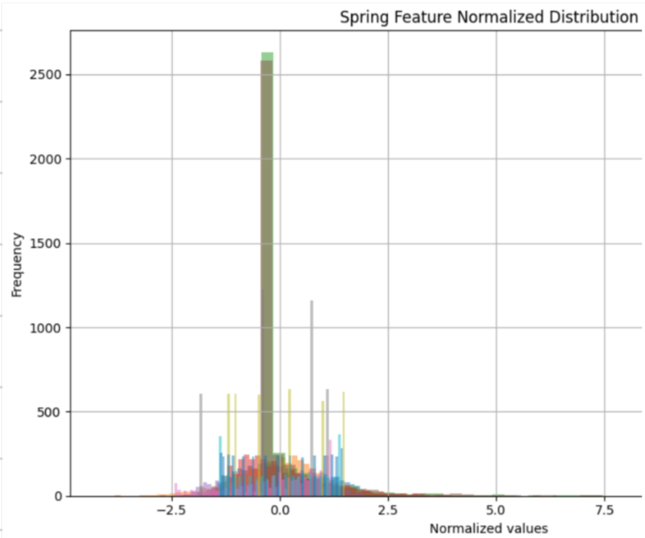


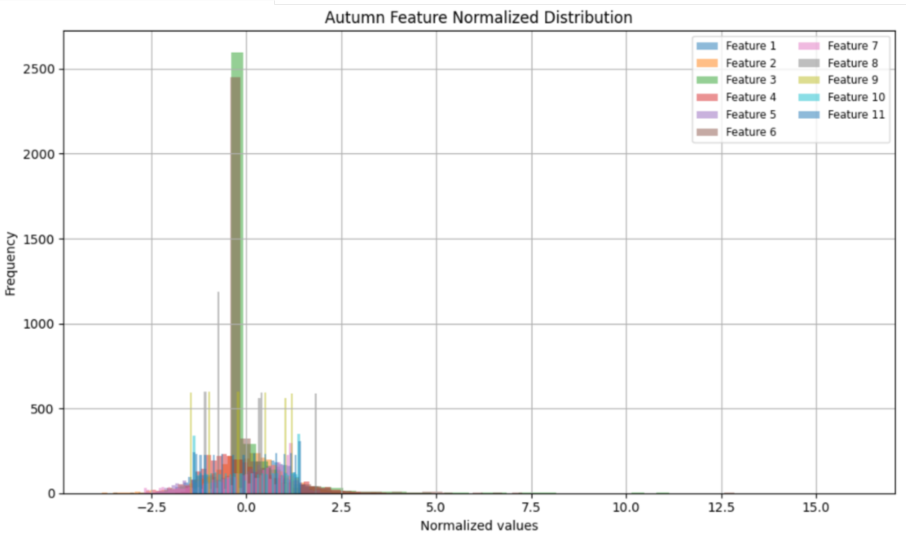
	SPRING	AUTUMN
RMSE	38.17 ±	73.93 −
R^2	0.35 ±	-0.01 −





Let's take a step back and analyze the distribution of data:





This data suits better "tanh" activation function.

Model 2.0:

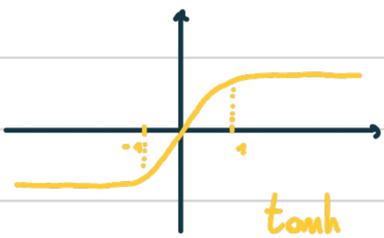
ACTIVATION FUN: tanh ; LOSS: MSE

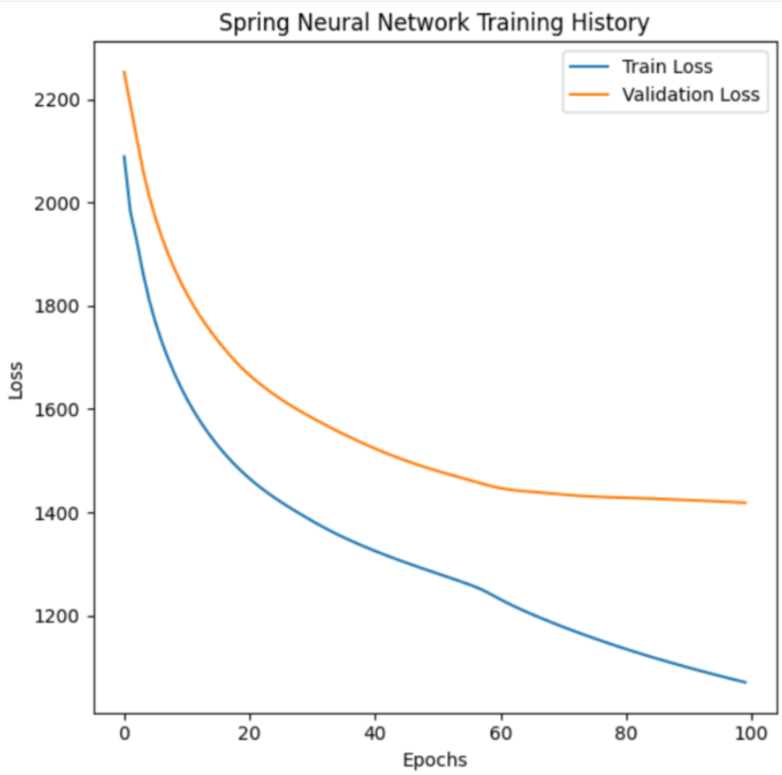
LEARNING RATE: ADAM, η = 0.001

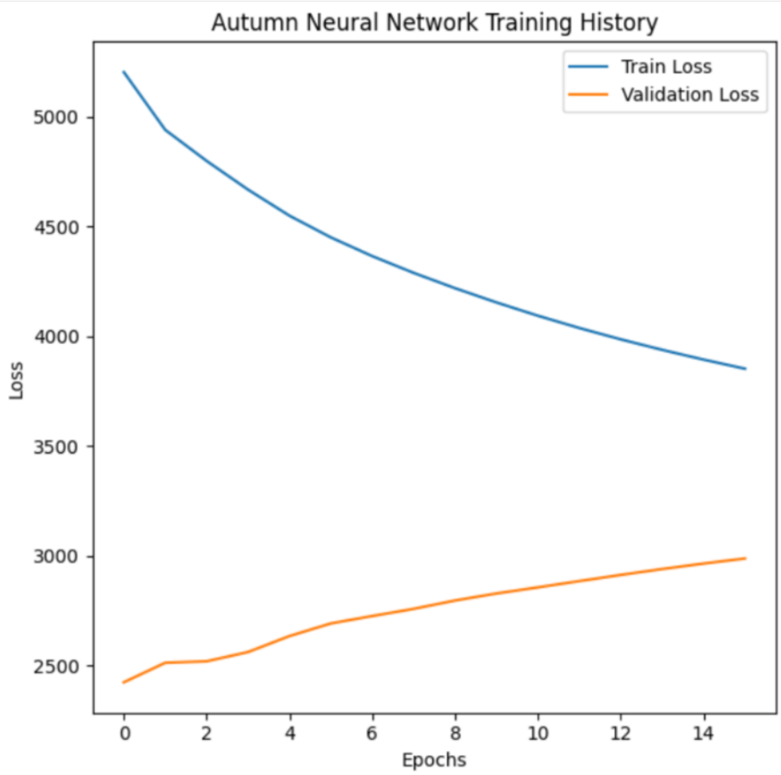
EPOCHS: 100

BATCH SIZE: 30

REGULARIZERS: μ = 0.001

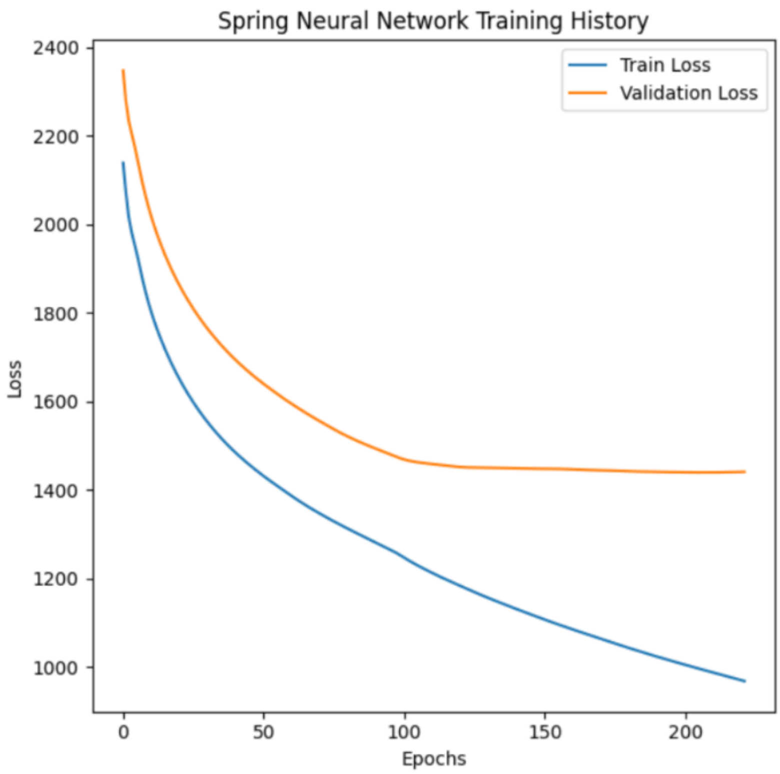


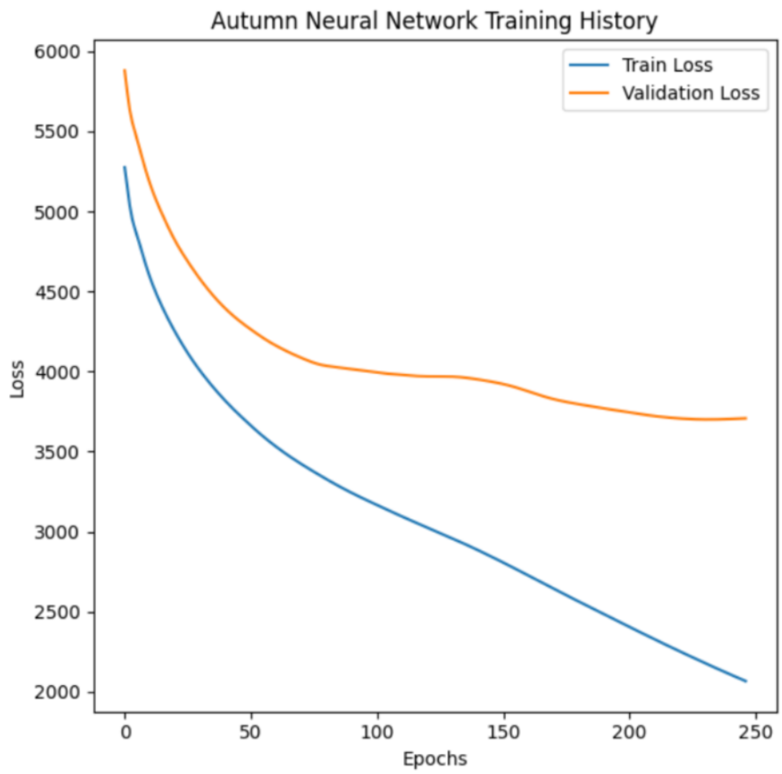




	SPRING	AUTUMN
RMSE	37.64 +	75.46 −
R^2	0.37 +	-0.05 −

Results aren't better yet, but the curve tells us to stabilize leaning.
Let's increase batch size From 30 to 90 to stabilize gradients and increase epochs.



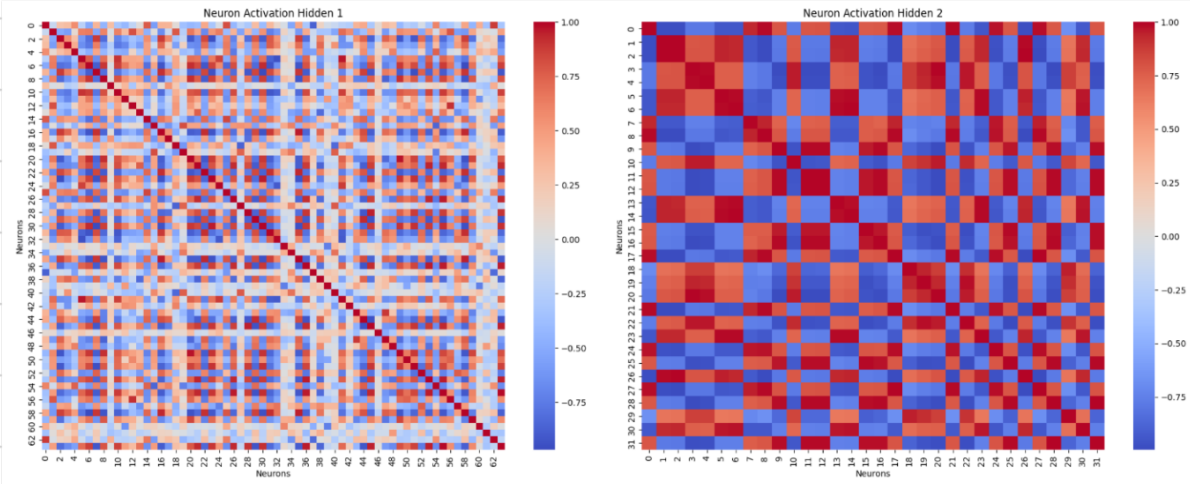


	SPRING	AUTUMN
RMSE	37.93 ±	60.8 +
R^2	0.36 ±	0.32 +

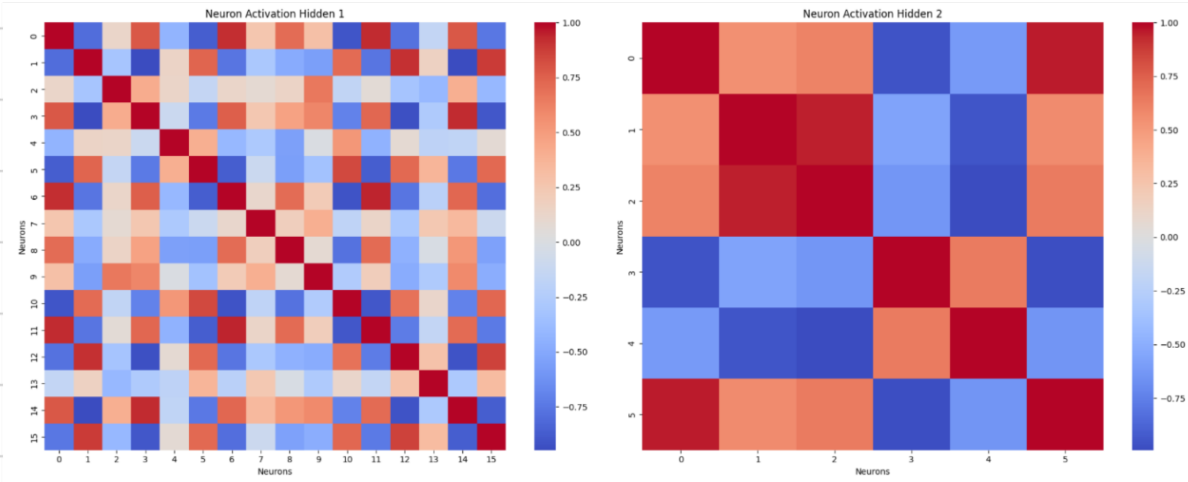
Finally better results!

Now I analyze Neurons activations to justify this "large" number of neurons :
Let's tweak Spring model numbers of neurons to reduce overfitting.

Before:
(64, 32)

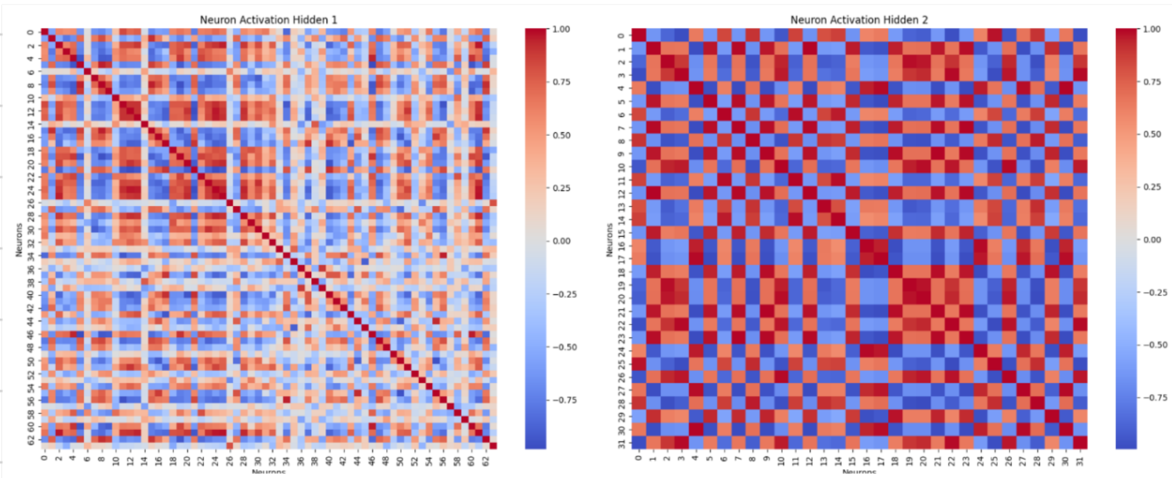


After:
(16, 6)

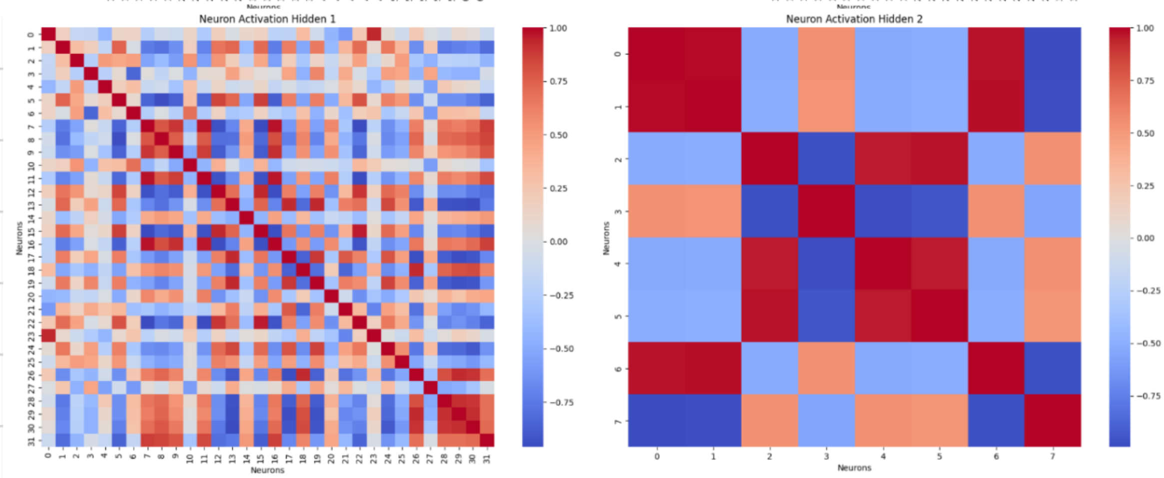


And now Autumn Model:

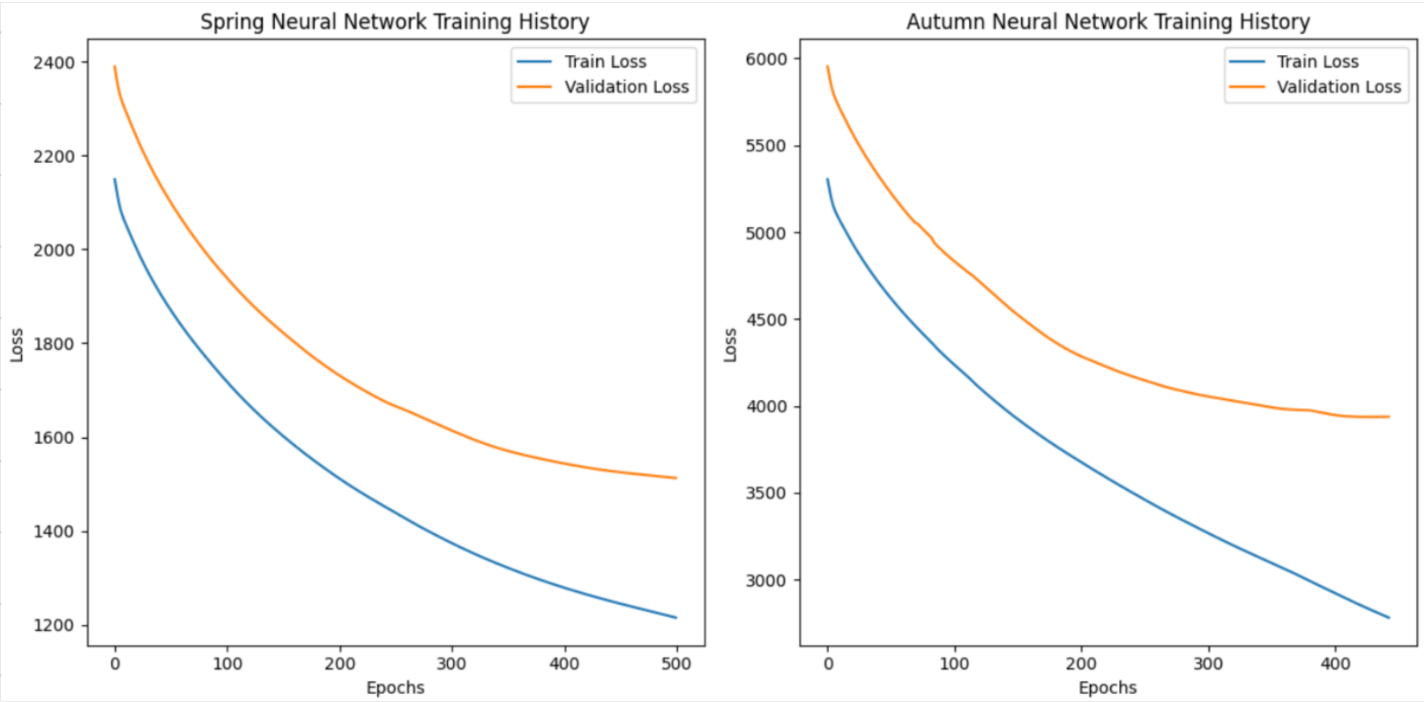
Before:
(64, 32)



After:
(32, 8)



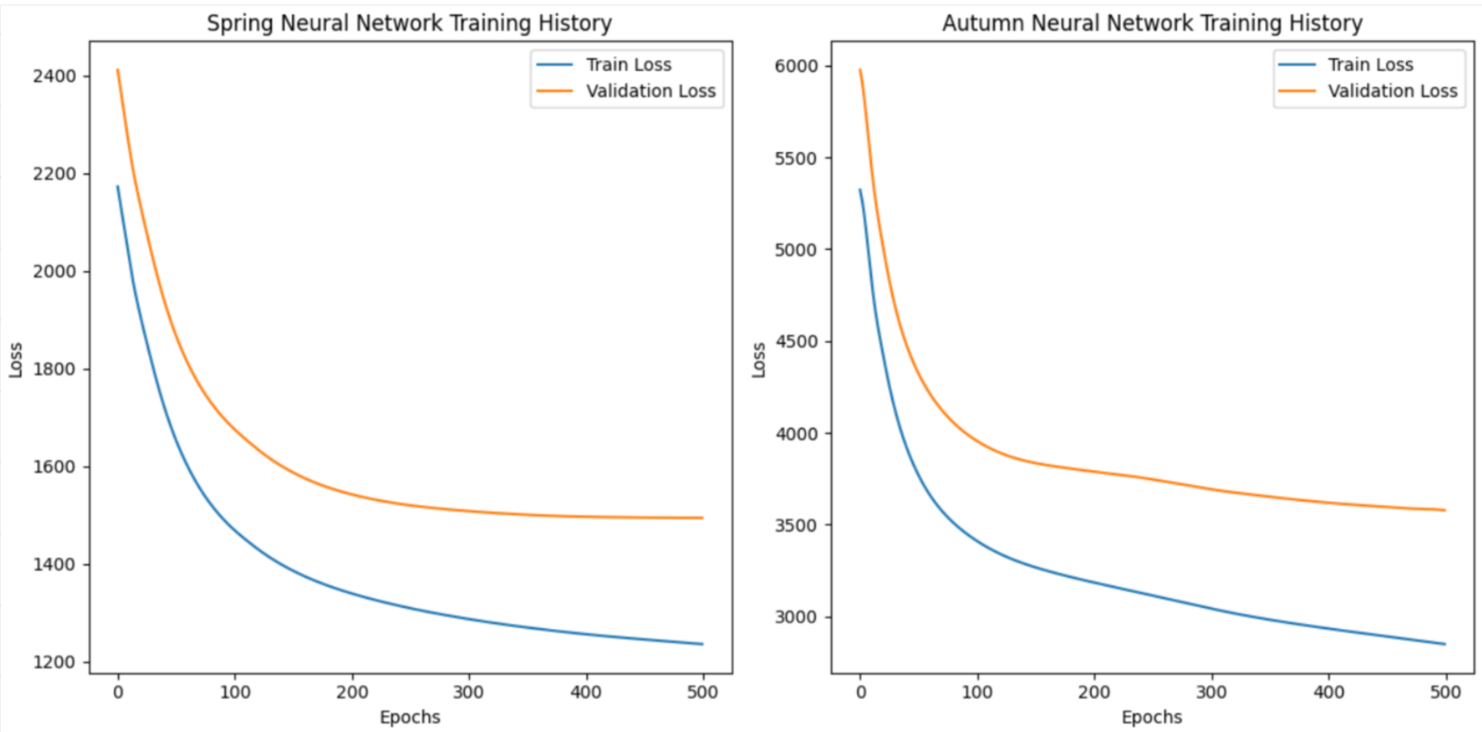
Let's check new performances:



	SPRING	AUTUMN
RMSE	38.89 ±	62.74 ±
R^2	0.33 ±	0.27 ±

But we still get overfitting at a specific point.
Do the two hidden layers help here?

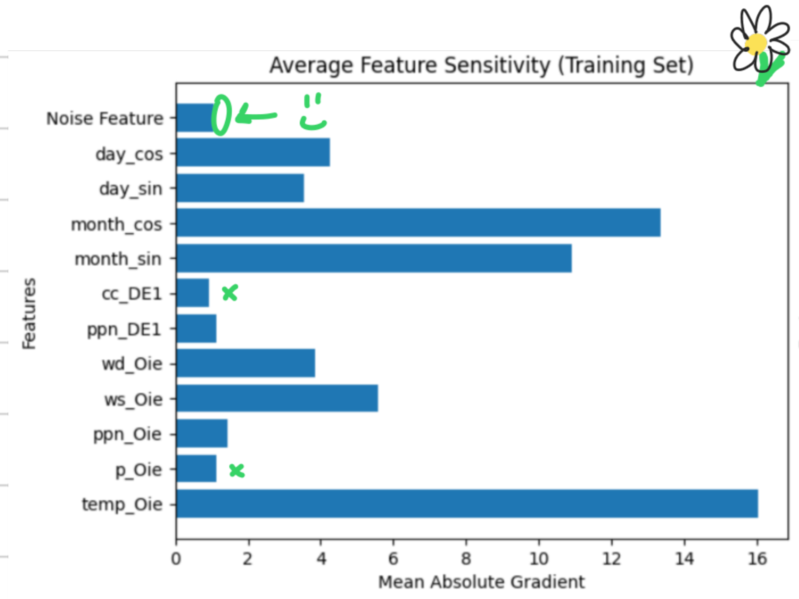
=> Let's remove "hidden 2" from both Spring and Autumn!

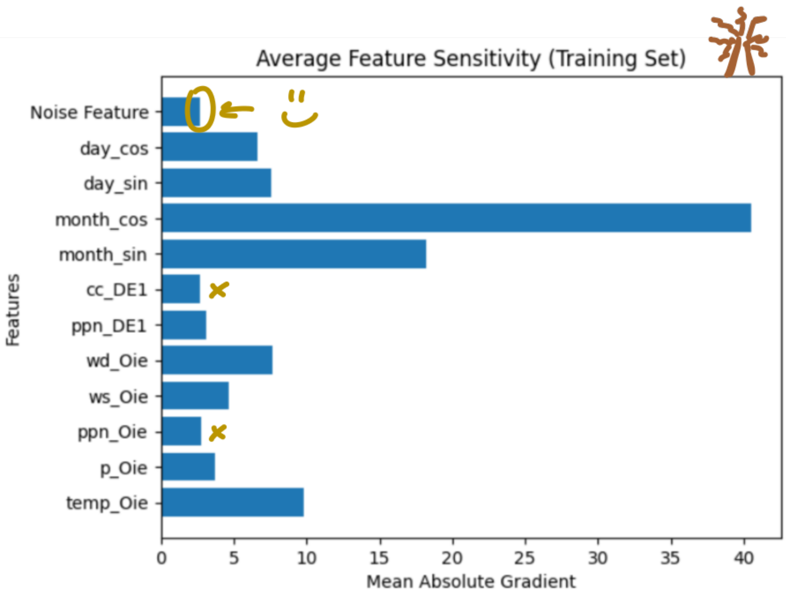


	SPRING	AUTUMN
RMSE	38.65 +	59.81 +
R^2	0.34 ±	0.34 +

· This is the best result we got so far.

What if I add some NOISE





· Let's do Ensemble to increase final performance!

Model 3.0:

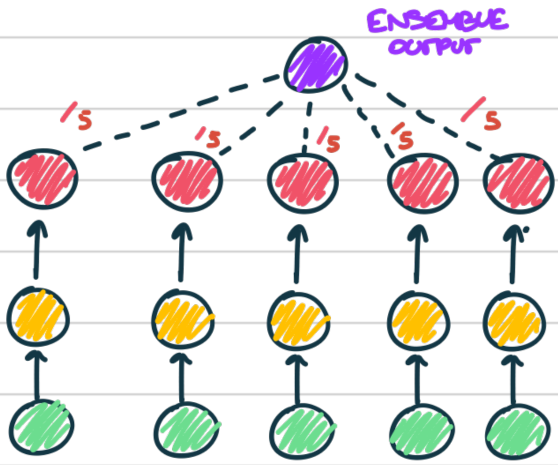
ACTIVATION FUN: tanh ; LOSS: MSE

LEARNING RATE: ADAM, η = 0.001

EPOCHS: 500

BATCH SIZE: 90

REGULARIZERS: μ = 0.001



SPRING: m = 16

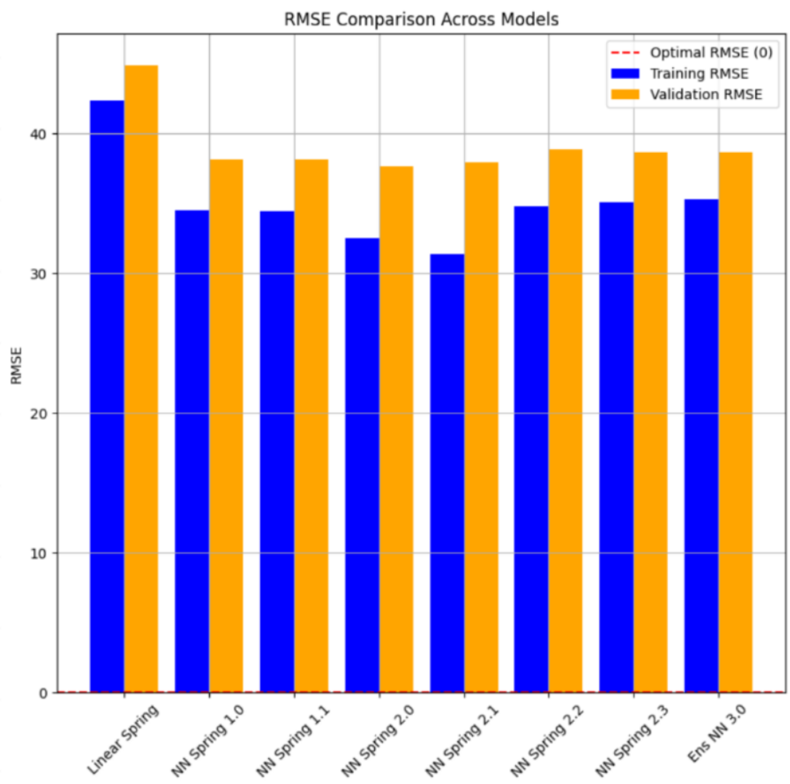
AUTUMN: m = 32

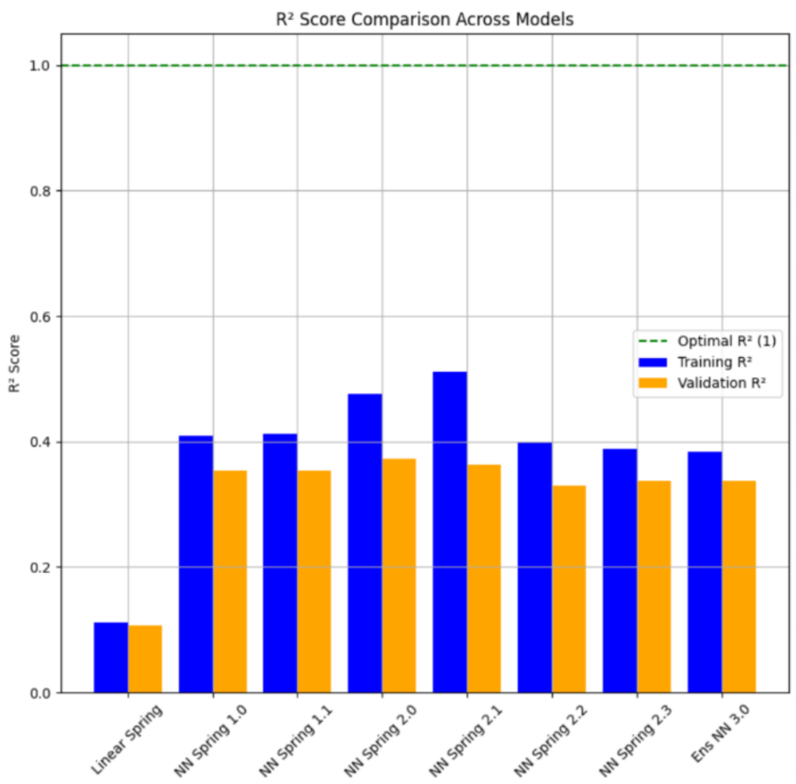
	SPRING	AUTUMN
RMSE	38.65 ±	59.94 ±
R^2	0.34 ±	0.34 ±

PERFORMANCE EVALUATION

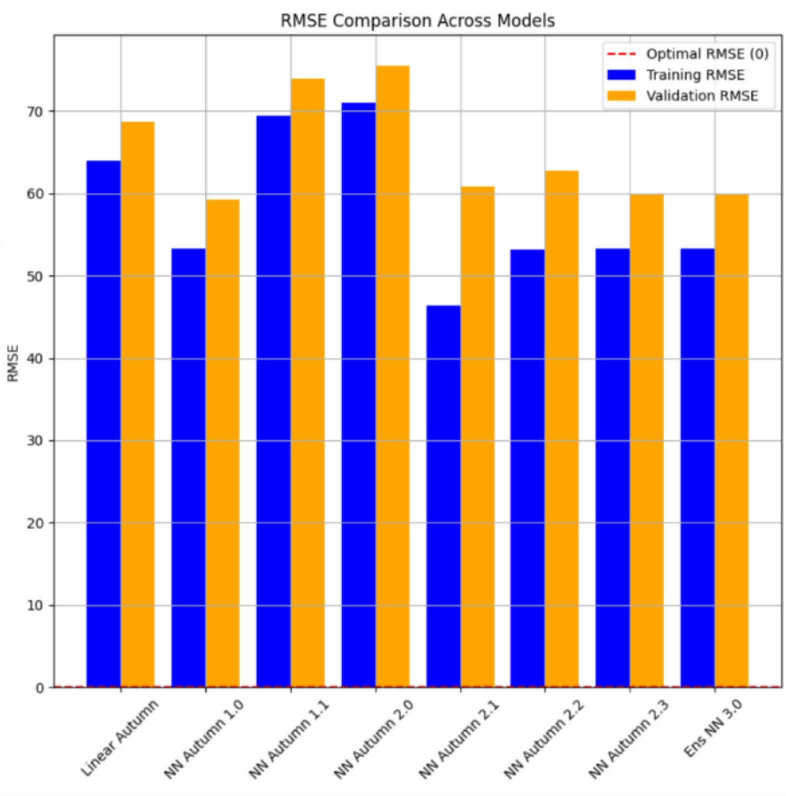
Let's look at the Final graph to draw conclusions:

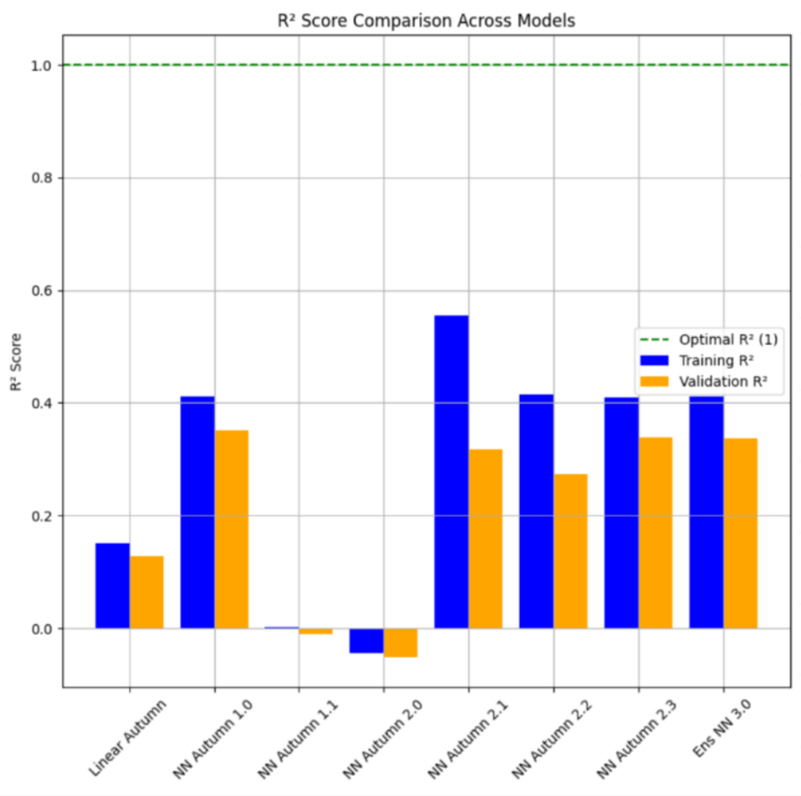
SPRING





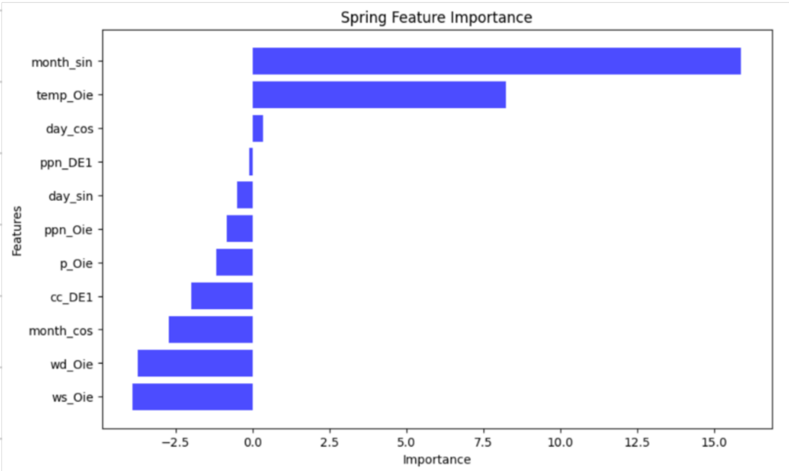
AUTUMN

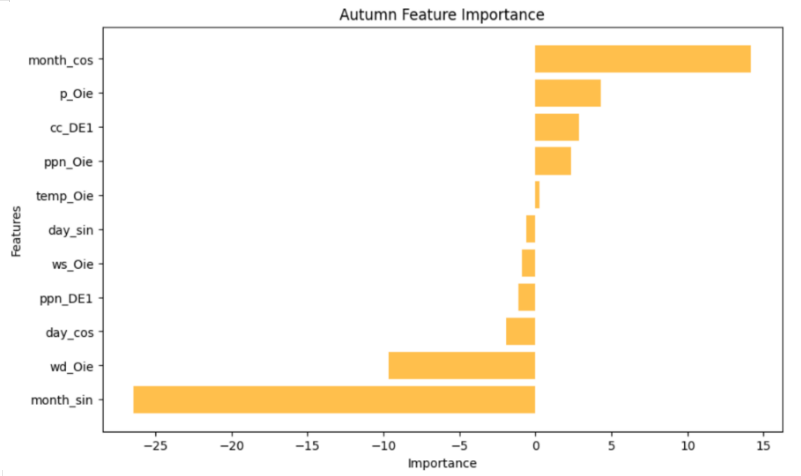




FEATURE IMPORTANCE

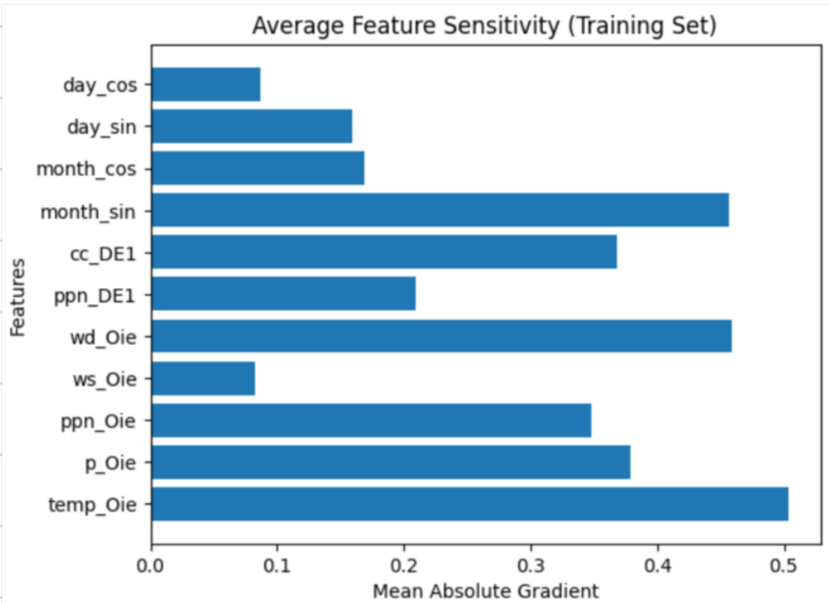
Linear Model Feature importance:



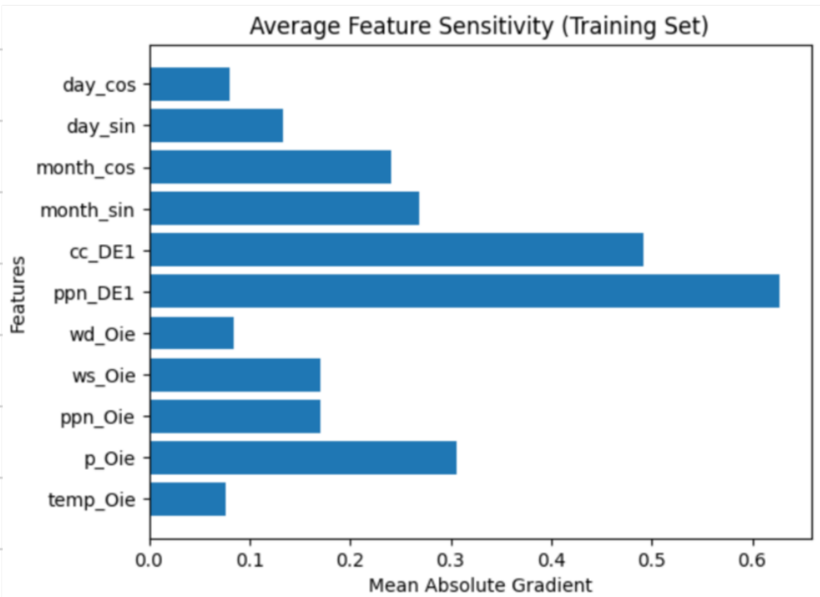


Ensemble 3.0 Neural Network Model:

Spring:



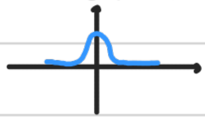
Autumn:



CONCLUSION

① We analyzed the dataset: ⇒ Correlation Matrices
⇒ Logical thinking

② We Modeled data: | 80% TRAIN | 20% TEST | ⇨ Normal distribution



③ We built a linear Model and multiple Neural Network Models.

WHAS IT USFULL? ⇨ YES!

We learned that bird migration data is highly NON-linear.

There is much more to learn, analyze and improve ☺