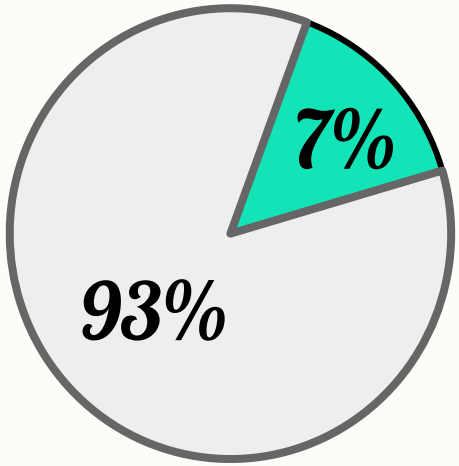
Diabetic Retinopathy

Diabetic Retinopathy

Why we should this Dataset



Thai population
69 Million

7% = 4.83 Million has diabete

50% of 7% don't know has diabete

Why should we fear?

Diabetic Retinopathy

What is diabete ?

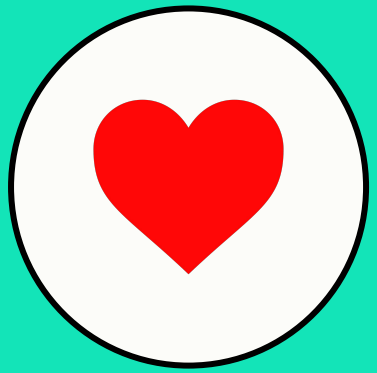
High blood sugar

INSULINE

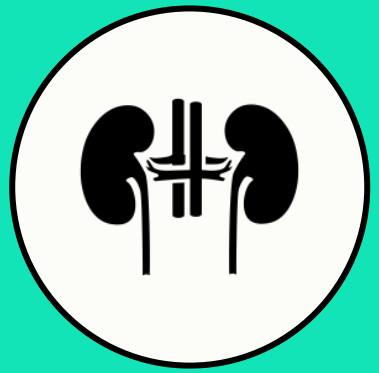
!

Diabetic Retinopathy

Complication Diseases



Heart Disease

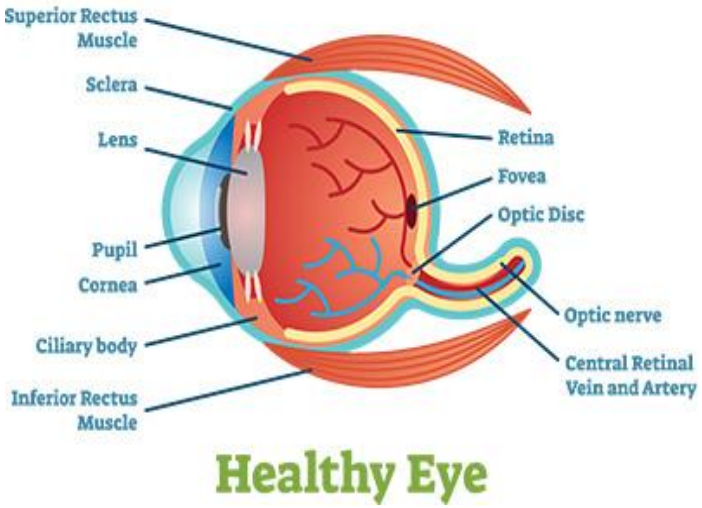


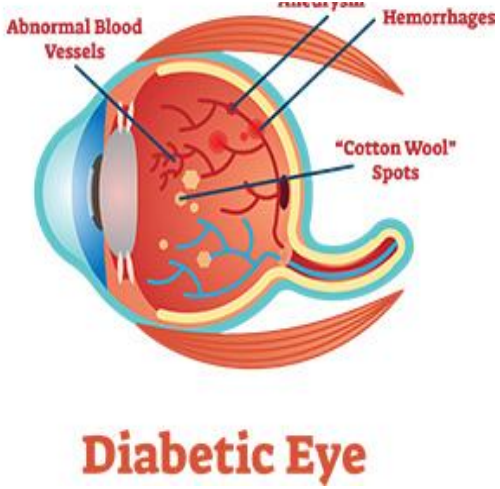
Kidney Disease

Diabetic Retinopathy

Diabetic Retinopathy

Diabetic Retinopathy





Cost

Case	Price (baht)
Basic check	2000-5,000
Advance check	5000-15000
Surgery	30,000-100,000++

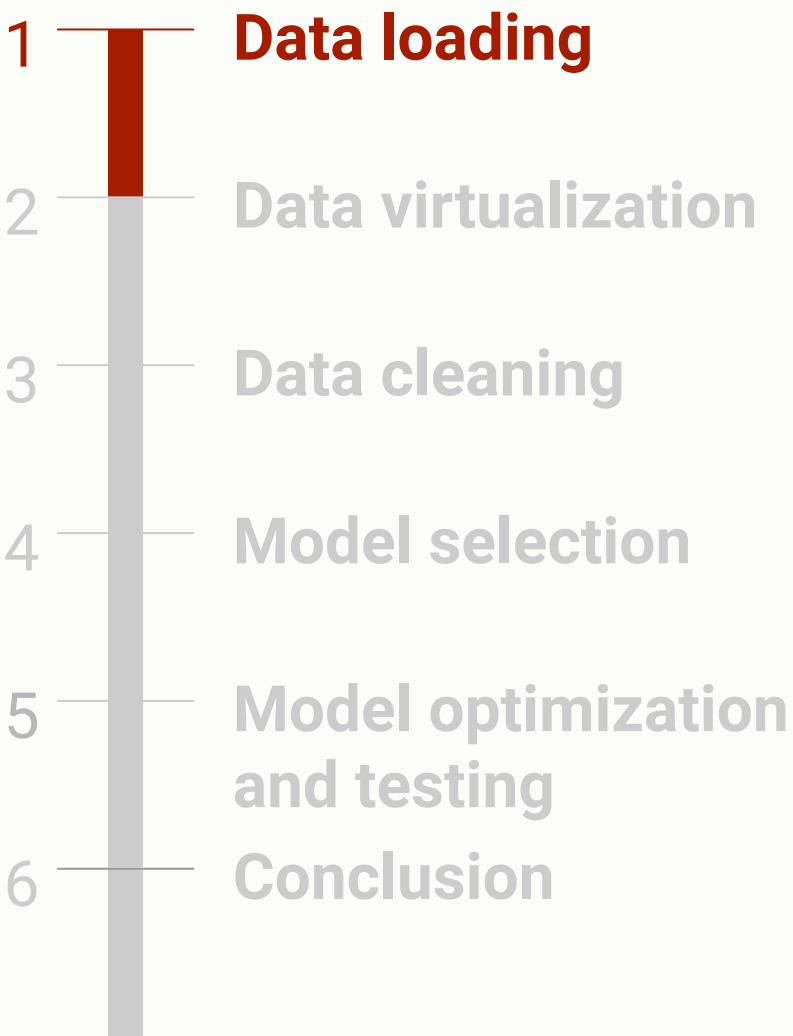
Diabetic Retinopathy

User

Doctor

Researcher

Patients


1
Data loading
2
Data virtualization
3
Data cleaning
4
Model selection
5
Model optimization and testing
6
Conclusion

Diabetic Retinopathy Dataset

- Size is 1,151 entries.
- There are 19 features.

	q	ps	nma.a	nma.b	nma.c	nma.d	nma.e	nma.f	nex.a	nex.b	nex.c	nex.d	nex.e	nex.g	nex.f	nex.h	dd	dm	amfm	class
0	1.0	1.0	22.0	22.0	22.0	19.0	18.0	14.0	49.895756	17.775994	5.270920	0.771761	0.018632	0.006864	0.003923	0.003923	0.486903	0.100025	1.0	b'0'
1	1.0	1.0	24.0	24.0	22.0	18.0	16.0	13.0	57.709936	23.799994	3.325423	0.234185	0.003903	0.003903	0.003903	0.003903	0.520908	0.144414	0.0	b'0'
2	1.0	1.0	62.0	60.0	59.0	54.0	47.0	33.0	55.831441	27.993933	12.687485	4.852282	1.393889	0.373252	0.041817	0.007744	0.530904	0.128548	0.0	b'1'
3	1.0	1.0	55.0	53.0	53.0	50.0	43.0	31.0	40.467228	18.445954	9.118901	3.079428	0.840261	0.272434	0.007653	0.001531	0.483284	0.114790	0.0	b'0'
4	1.0	1.0	44.0	44.0	44.0	41.0	39.0	27.0	18.026254	8.570709	0.410381	0.000000	0.000000	0.000000	0.000000	0.000000	0.475935	0.123572	0.0	b'1'

Diabetic Retinopathy

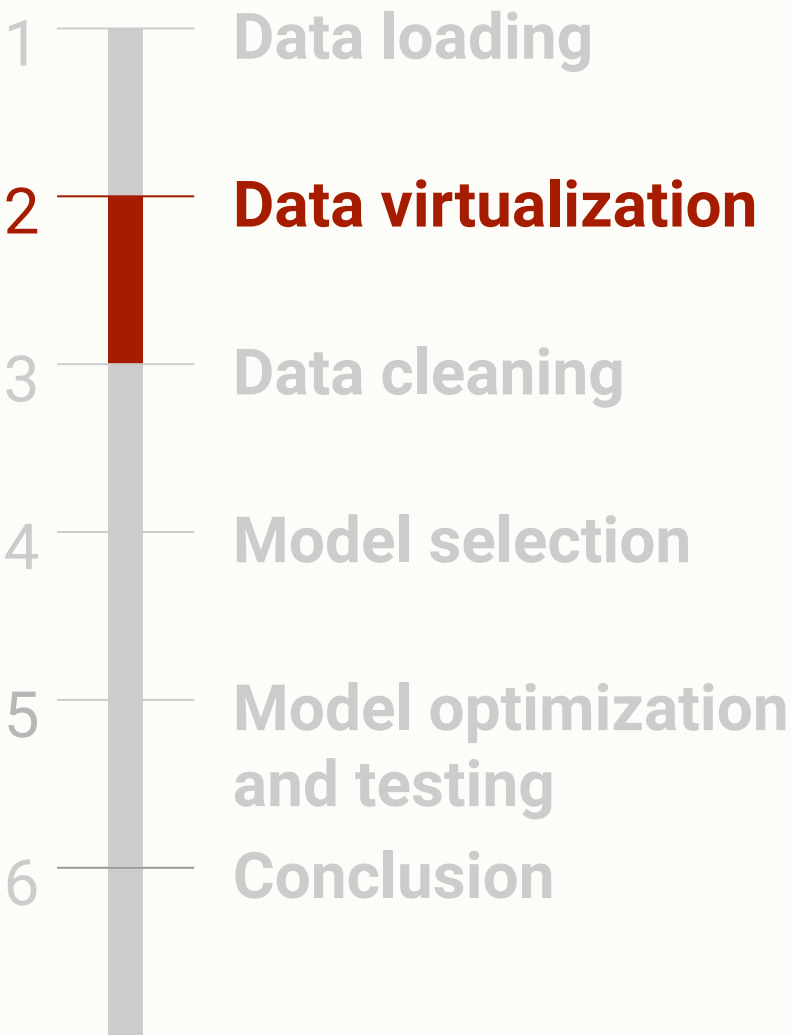
Features

- "q" = The binary result of quality assessment.
- "ps" = The binary result of pre-screening.
- "nma.a" - "nma.f" = Number of Microaneurysms found at the confidence levels alpha from 0.5 to 1 repeatedly.

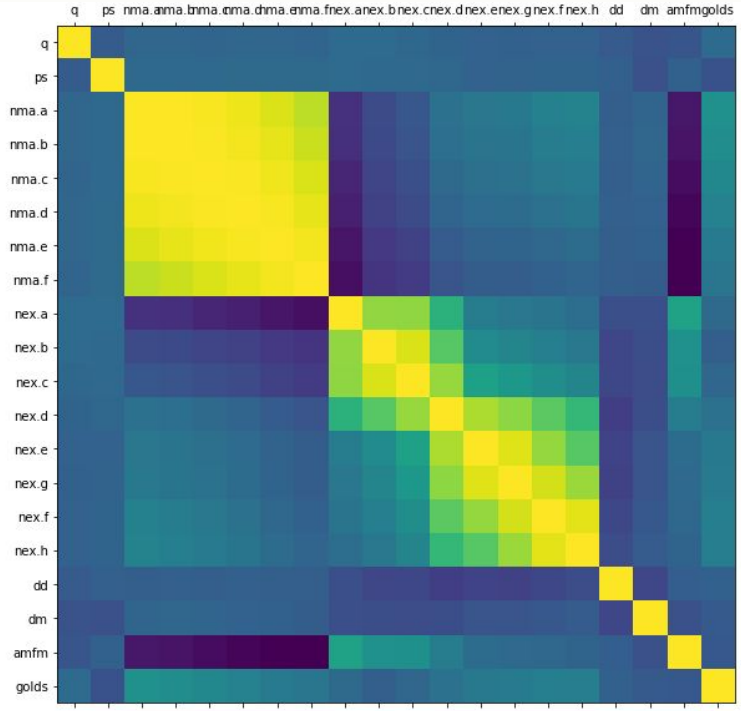
Diabetic Retinopathy

Features

- "nex.a" - "nex.h" = Number of Exudates found at the confidence levels alpha from 0.3 to 1 repeatedly.
- "dd" = The euclidean distance of the center of the macula and the center of the optic disc.
- "dm" = The diameter of the optic disc.
- "amfm" = The binary result of the AM/FM-based classification.


1
Data loading
2
Data virtualization
3
Data cleaning
4
Model selection
5
Model optimization
and testing
6
Conclusion

Diabetic Retinopathy
Dataset Behavior



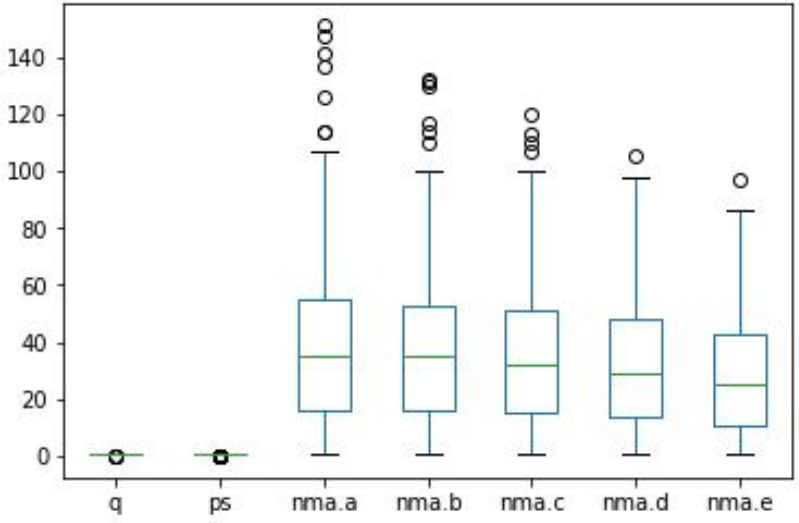
Covariance Matrix

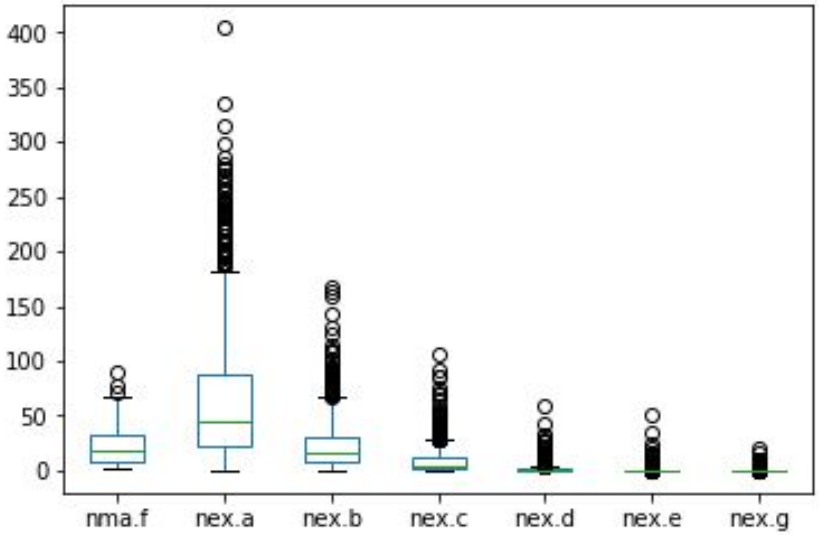
- nma.a - nma.f are correlated.
- same as nex.a - nex.c
- same as nex.d - nex.h
- There are some relation b/w golds, nma.a, nex.h

Diabetic Retinopathy

Dataset Behavior

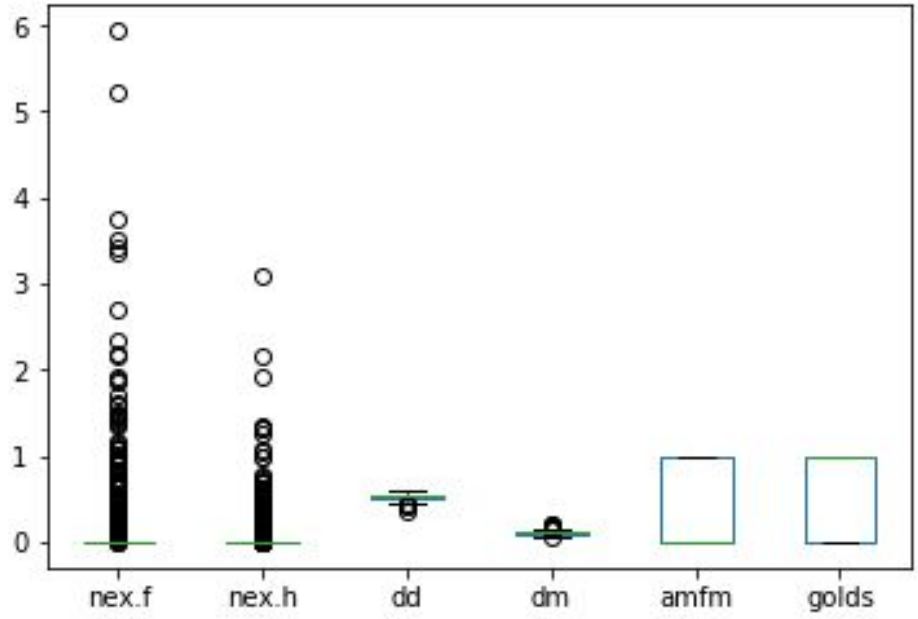
Outliers :





Diabetic Retinopathy
Dataset Behavior

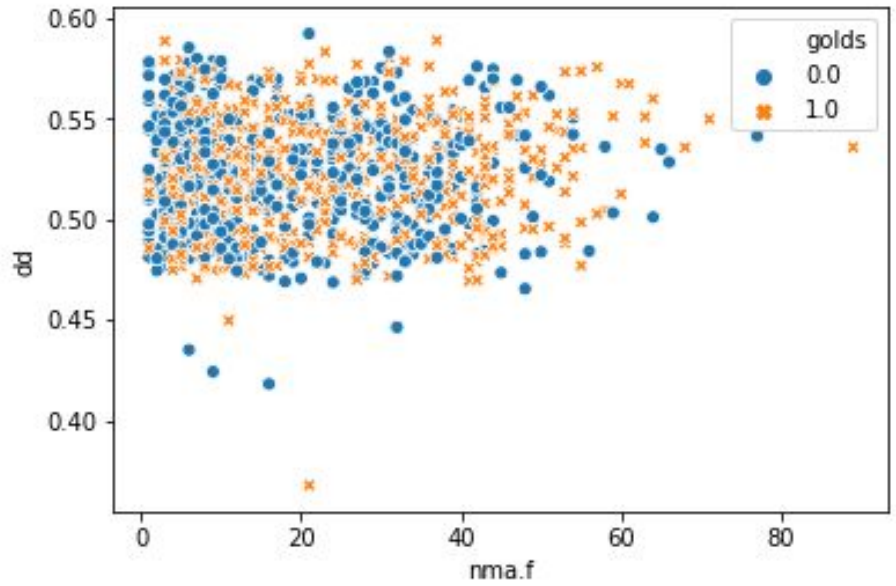
Outliers 2:

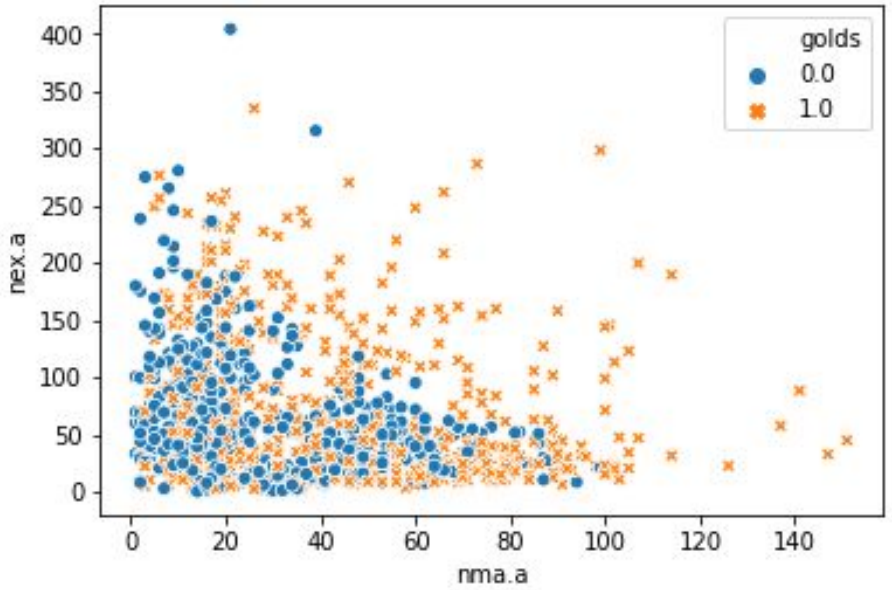


Diabetic Retinopathy

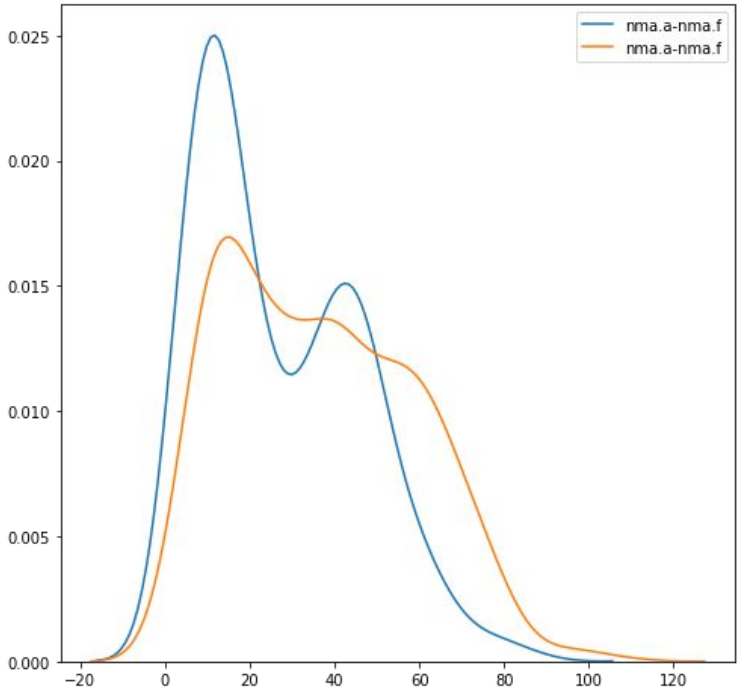
Dataset Behavior

- Data complexity is high, difficult to classify.

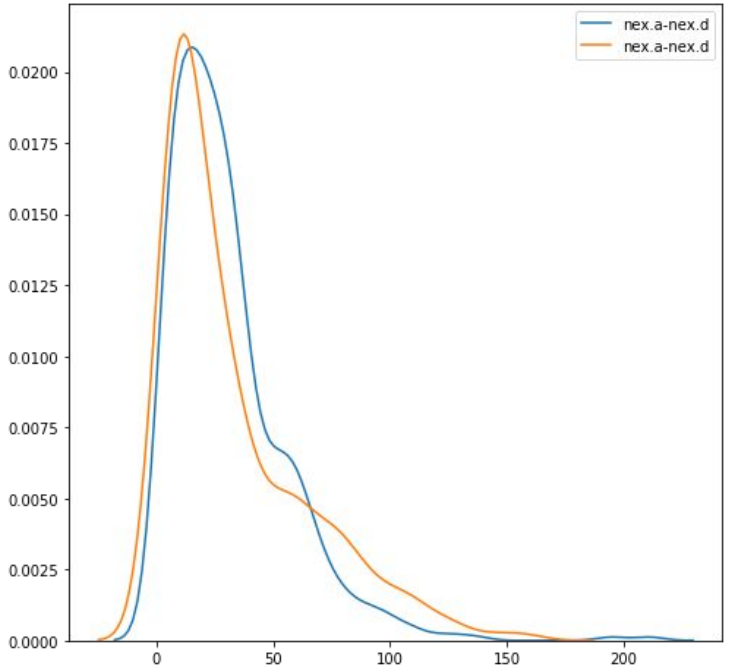




Diabetic Retinopathy
Dataset Behavior



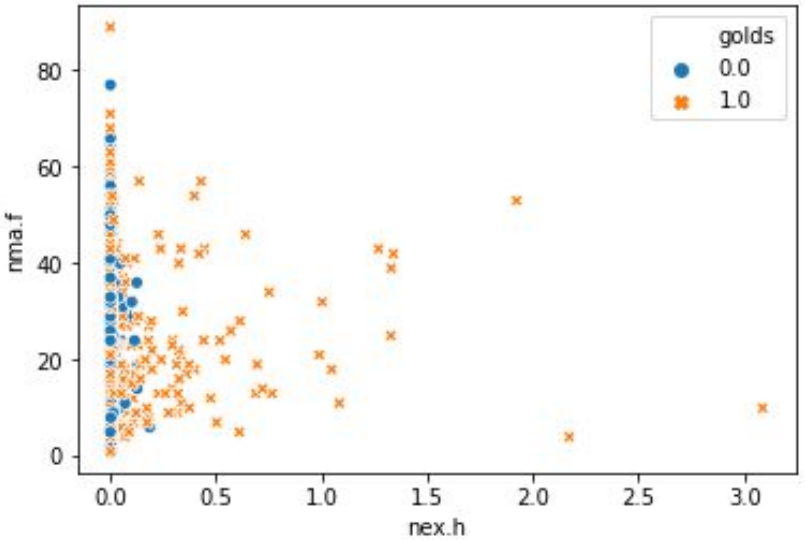
nma Positive result & Negative result.

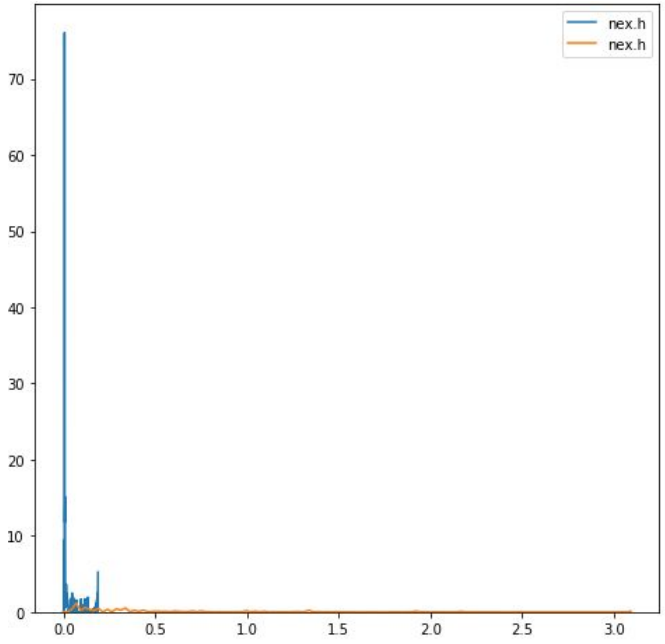


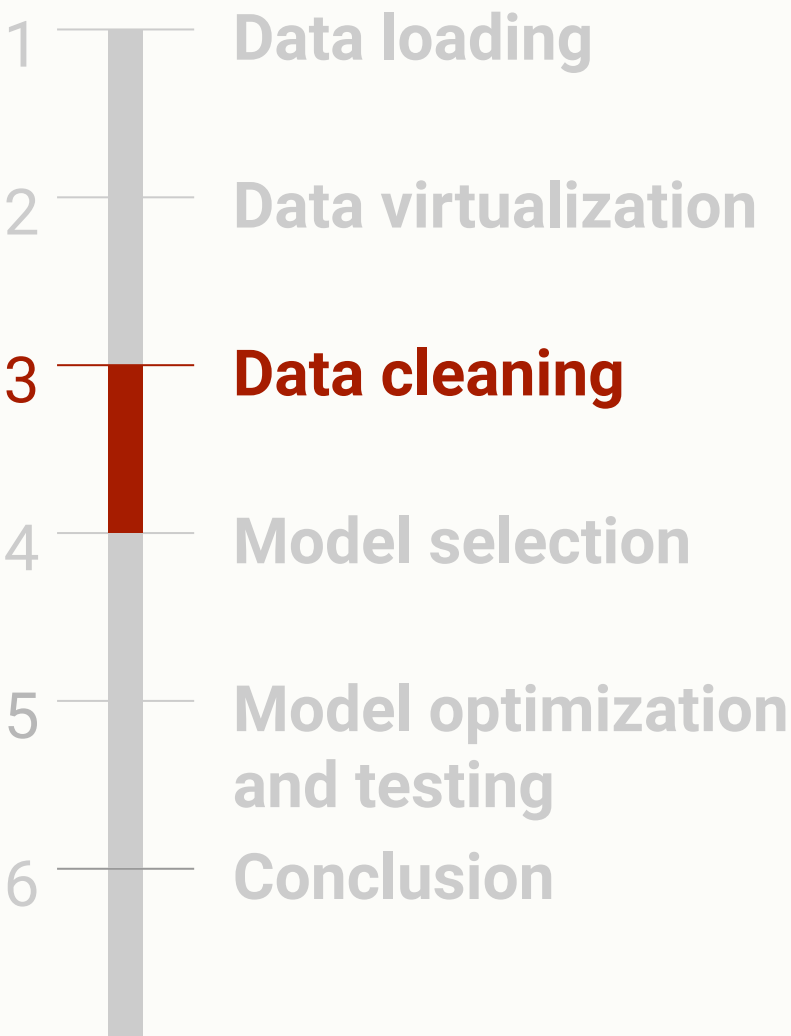
nex Positive result & Negative result

Diabetic Retinopathy
Dataset Behavior

- But, there are some feature still look okay to classify (e.g. nex.f, nex.h)






1
Data loading
2
Data virtualization
3
Data cleaning
4
Model selection
5
Model optimization
and testing
6
Conclusion

Diabetic Retinopathy

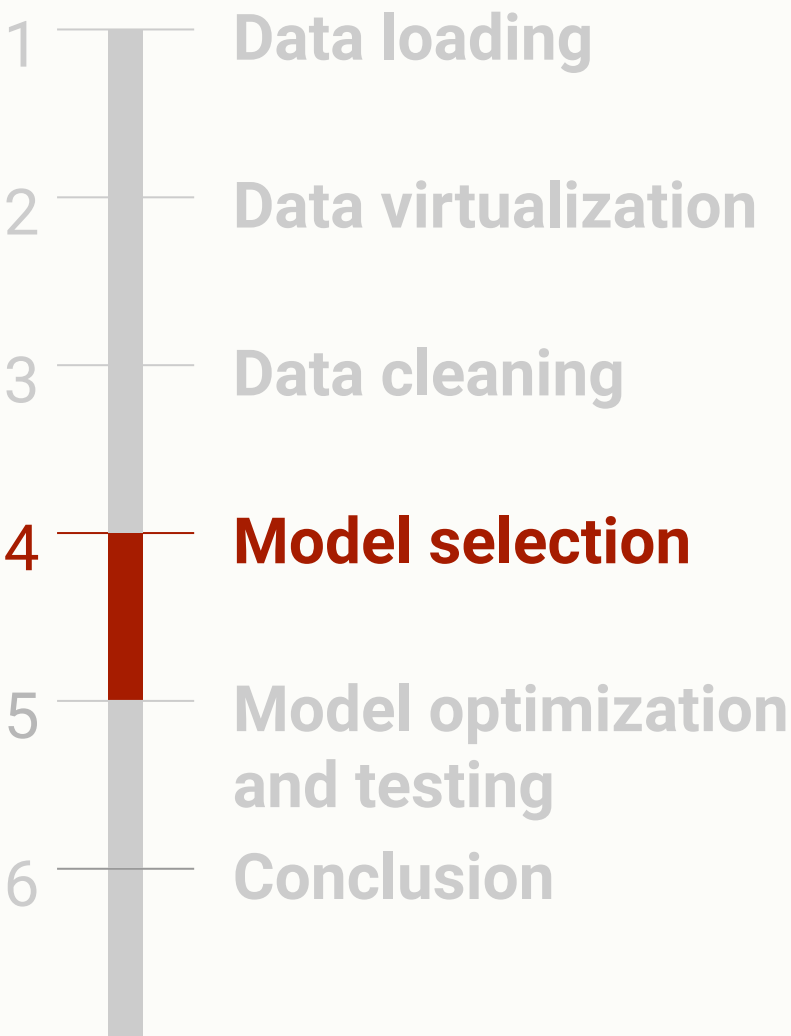
Cleansing Data

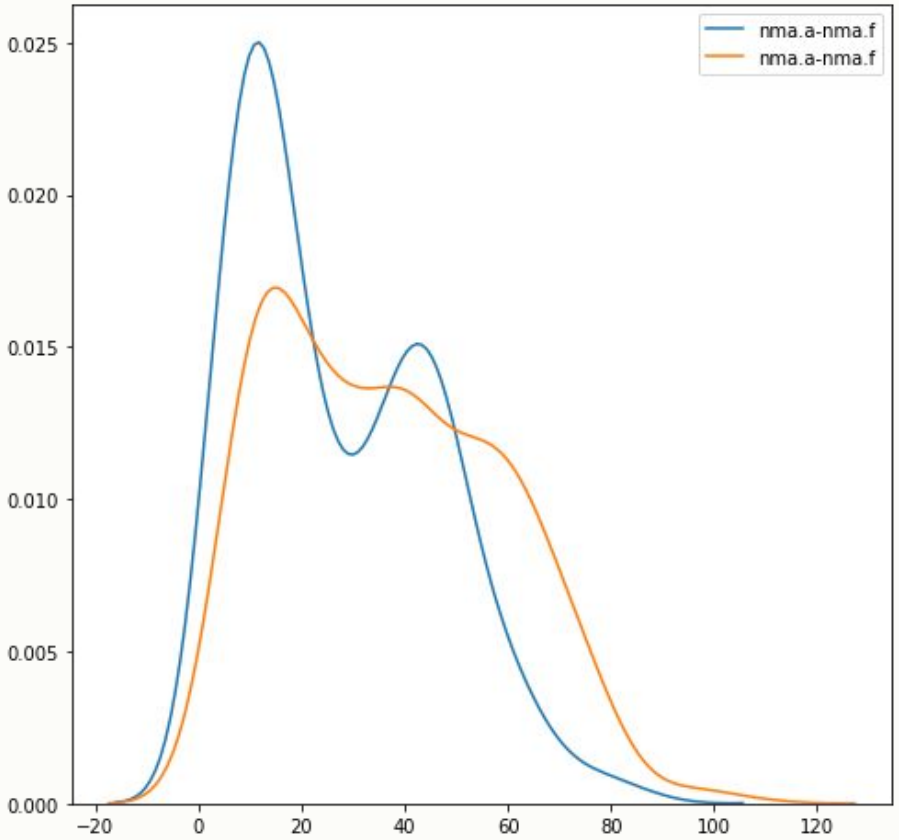
- We have noticed that "q" has mean = 0.996.
- Therefore, there are a few bad quality data.
- So, we remove all bad quality data.
- And, delete "q" feature.

Diabetic Retinopathy

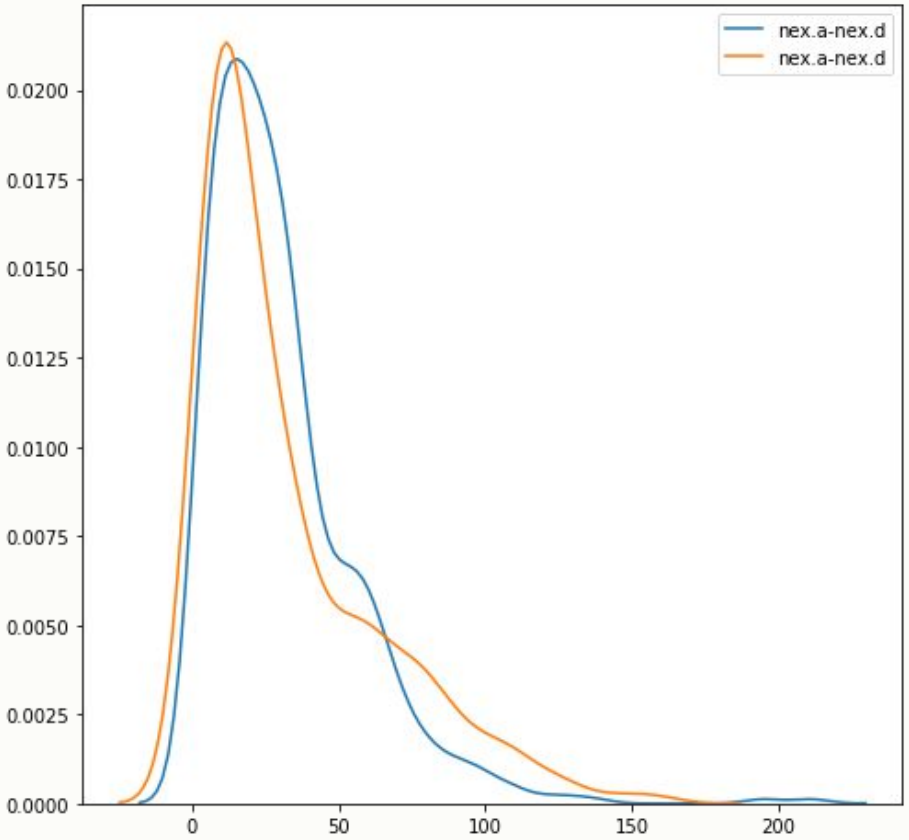
Cleansing Data

- For outliers, we have done winsorizing.
- The reason, eventually there are so much outliers, but they are in the features that have high confidence.


1
Data loading
2
Data virtualization
3
Data cleaning
4
Model selection
5
Model optimization and testing
6
Conclusion



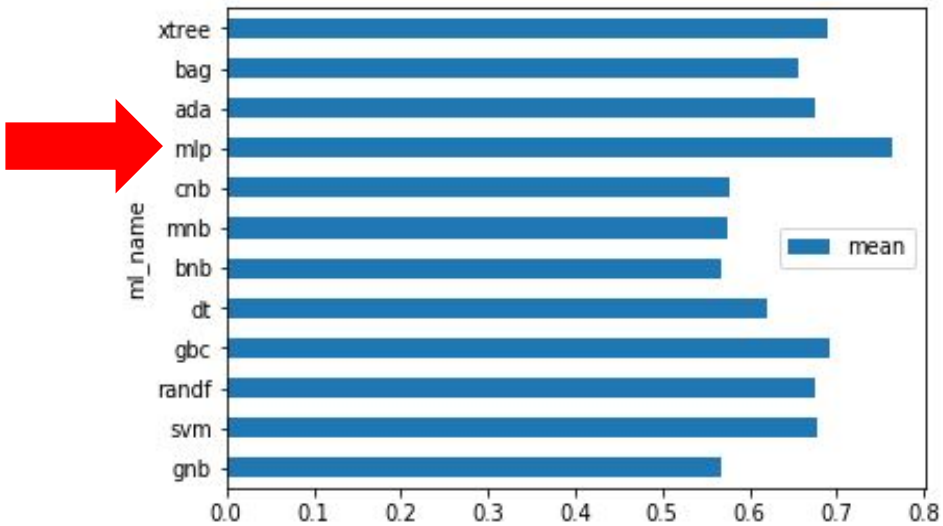
nma Positive result & Negative result.



nex Positive result & Negative result

Diabetic Retinopathy
Model Selection

- We select Multi-Layer Perceptron.
  - After we test so many MLs, we found that MLP is the best.

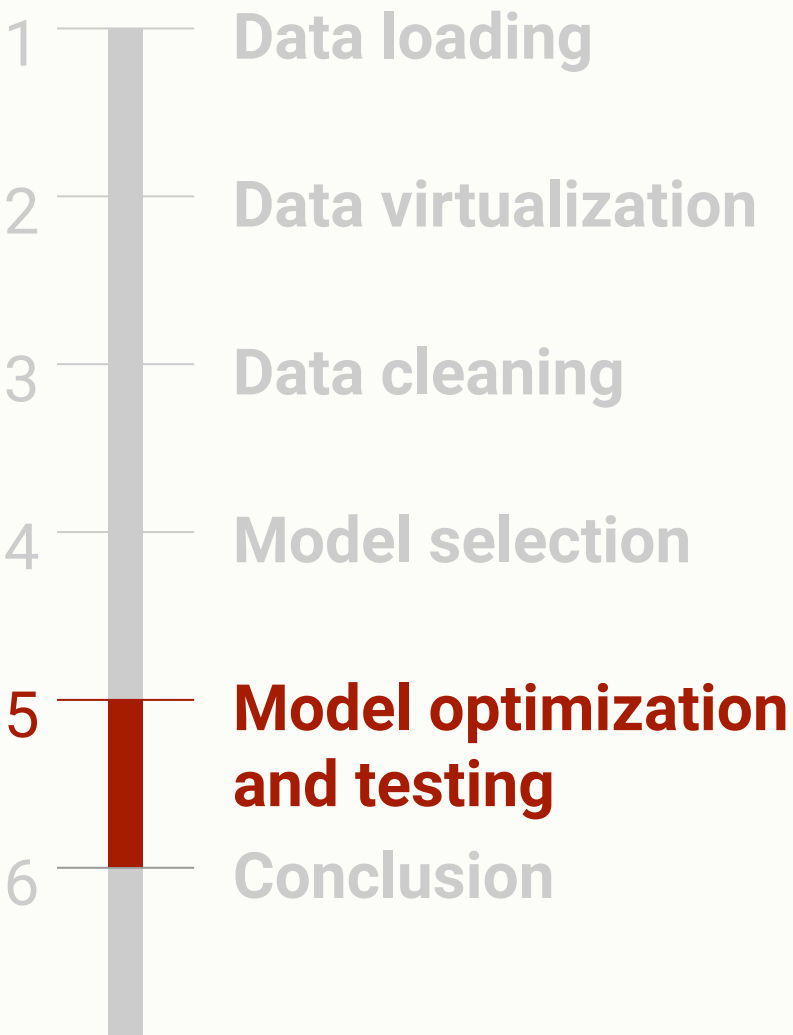


Diabetic Retinopathy

Model Benefits

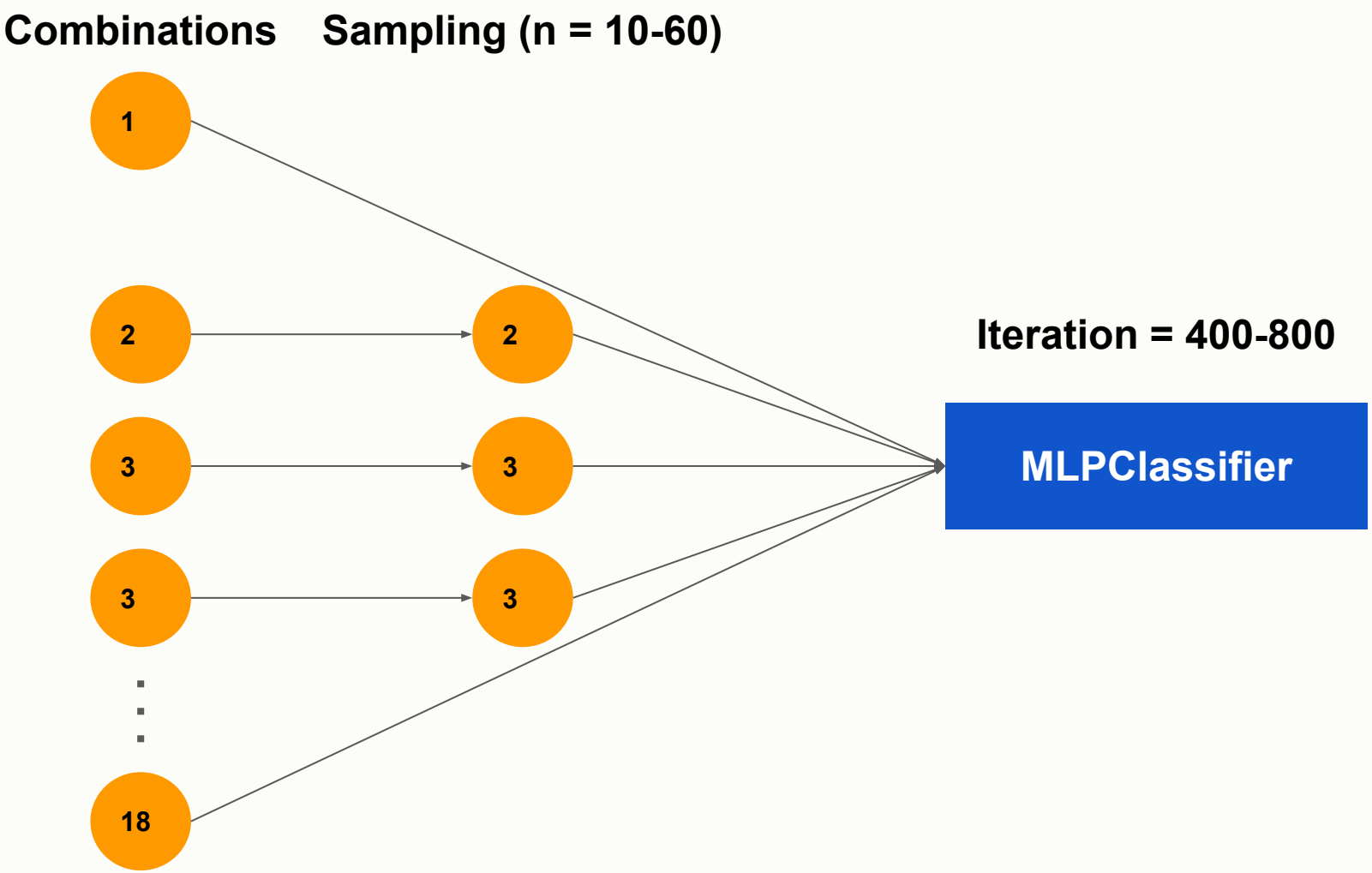
Advantage of MLP-Classification.

- It is good for classification prediction problems .
- Can perform multiple boundaries to separate two classes, which is good for our dataset.

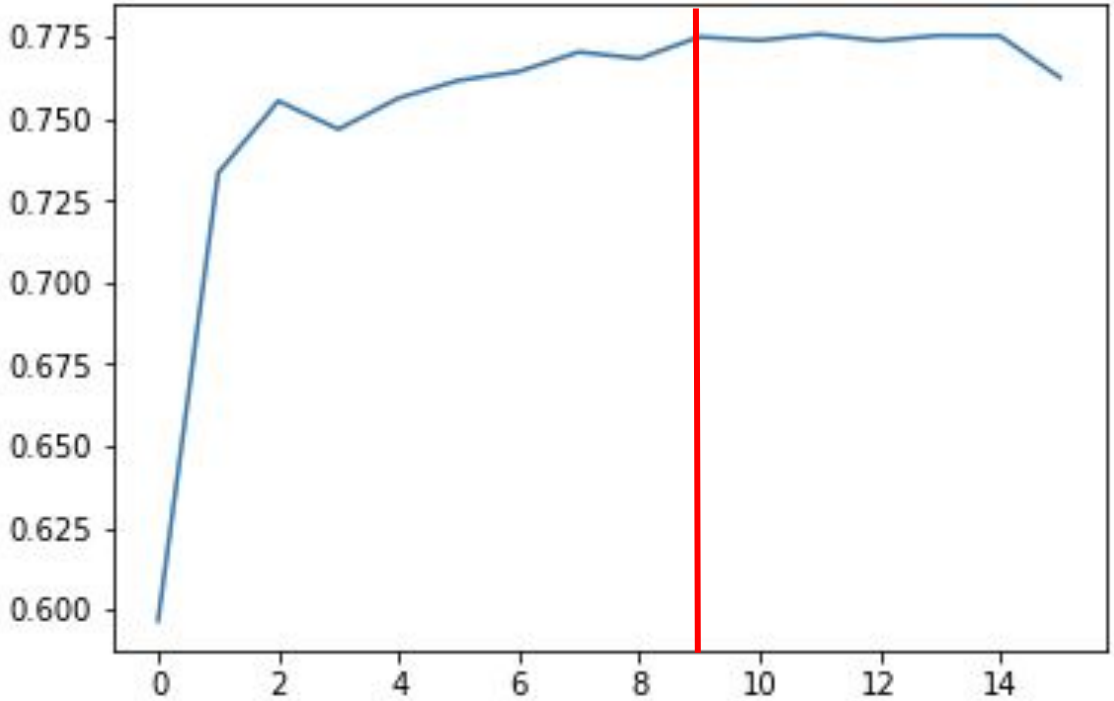

1
Data loading
2
Data virtualization
3
Data cleaning
4
Model selection
5
Model optimization and testing
6
Conclusion

F1 Scores

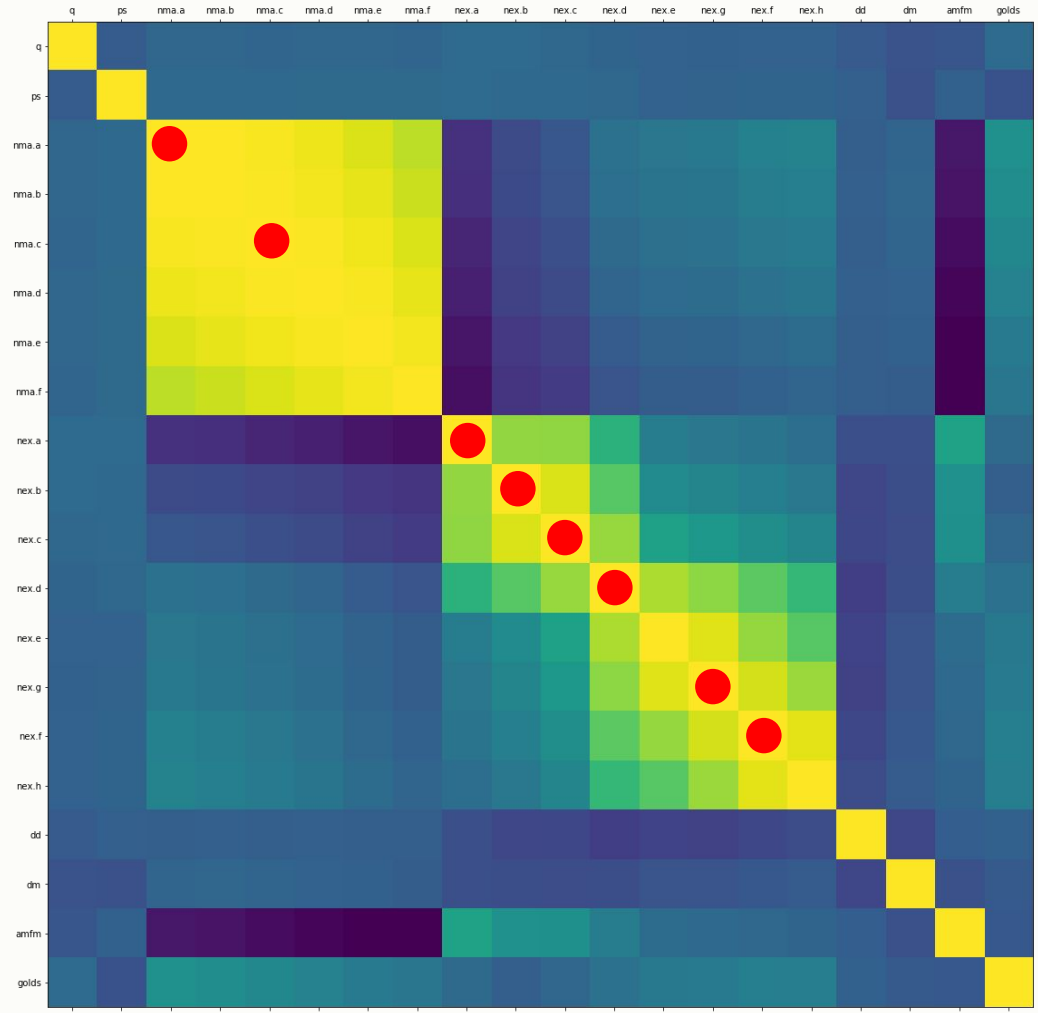
All features (18)	Feature important (10)	Filtered features (8)
0.75796	0.74496	0.76348


Combinations
Sampling (n = 10-60)
1
2
2
3
3
3
3
18
Iteration = 400-800
MLPClassifier

F1 Scores from “cross_validation_scores” at cv=4



	precision	recall	f1-score	support
0.0	0.72	0.83	0.77	136
1.0	0.82	0.70	0.76	150
micro avg	0.76	0.76	0.76	286
macro avg	0.77	0.77	0.76	286
weighted avg	0.77	0.76	0.76	286



nma.a
nma.c
nex.a
nex.b
nex.c
nex.d
nex.g
nex.f

Microan eurysms

Exudates

1 Data loading

2 Data virtualization

3 Data cleaning

4 Model selection

5 Model optimization and testing

6 **Conclusion**

Diabetic Retinopathy

Conclusion

- Exudates is a significant feature to tells if cataract formed by being diabetic.
- By using Brute-force combination feature extraction, we can extract critical features to find insight.

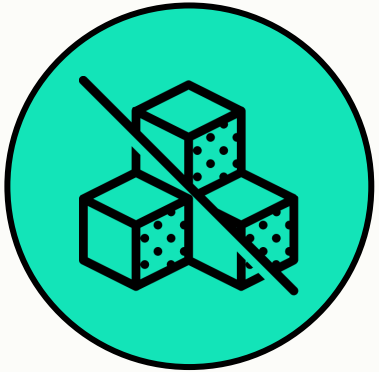
Diabetic Retinopathy

Impact

- This model can be used to tell if cataract patient is risk of getting complication diseases.

Diabetic Retinopathy

Solution



Low the sugar

Exercise

Carbohydrate Control

Diabetic Retinopathy

Carbohydrate Facts

- Main Energy of your body
- Rice is the main Carbohydrate of thai people
- Compare the another main carbohydrate

Rice > Noodle > Bread

Rice has the most carbohydrate in (100g)

Diabetic Retinopathy

Select the right rice

- Best choice : Rice (กข 43) has only 50% carbohydrate

- Another choice : Riceberry,Brown Rice has high fiber help in control cholesterol