ATMS 597
Project 5
Precipitation Classification: Supervised Machine Learning

Group F – Muskegon, Michigan

Xinchang Li (Cathy)

Max Grover

Piyush Garg

ASOS 5-min Data

14840KMKG
MKG20000131234511501/31/00
23:45:31 5-MIN KMKG 010445Z
30006KT 6SM -SN SCT018 OVC023
M02/M06 A2989 660 78 -1200 310

date_time	station_id	latitude	longitude	elevation	date_time.1	wind_direction	wind_speed	current_wx1	current_wx2	current_wx3
2000-01-01 00:00:00	KMKG	43.17	-86.23	191	2000-01-01 00:00:00	270.0	7.0	-SN	NaN	NaN
2000-01-01 00:05:00	KMKG	43.17	-86.23	191	2000-01-01 00:05:00	280.0	10.0	-SN	NaN	NaN
2000-01-01 00:10:00	KMKG	43.17	-86.23	191	2000-01-01 00:10:00	280.0	11.0	-SN	NaN	NaN
2000-01-01 00:15:00	KMKG	43.17	-86.23	191	2000-01-01 00:15:00	290.0	7.0	-SN	NaN	NaN
2000-01-01 00:20:00	KMKG	43.17	-86.23	191	2000-01-01 00:20:00	280.0	9.0	-SN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...
2020-03-30 18:25:00	KMKG	43.17	-86.23	191	2020-03-30 18:25:00	330.0	12.0	-RA	NaN	NaN
2020-03-30 18:30:00	KMKG	43.17	-86.23	191	2020-03-30 18:30:00	320.0	13.0	-RA	NaN	NaN
2020-03-30 18:35:00	KMKG	43.17	-86.23	191	2020-03-30 18:35:00	340.0	12.0	-RA	NaN	NaN
2020-03-30 18:45:00	KMKG	43.17	-86.23	191	2020-03-30 18:45:00	320.0	14.0	-RA	NaN	NaN
2020-03-30 18:50:00	KMKG	43.17	-86.23	191	2020-03-30 18:50:00	320.0	16.0	-RA	NaN	NaN

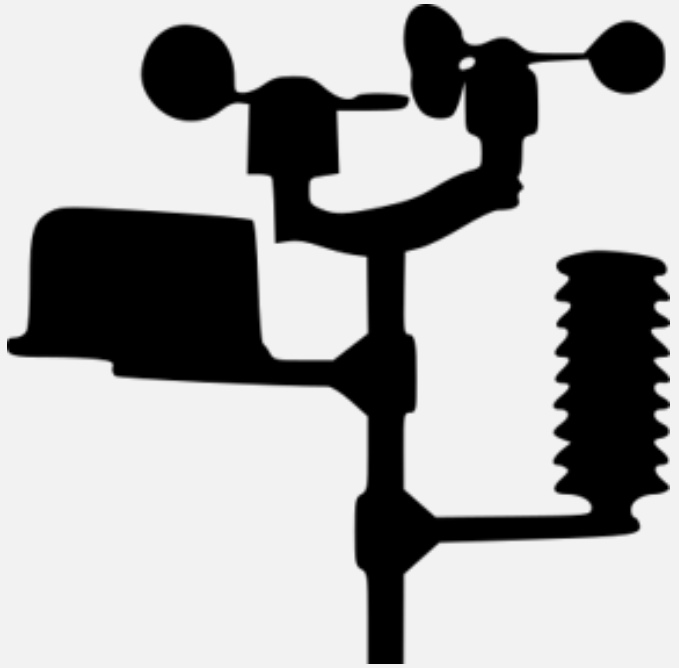
Geophysical Fields Used

Intrinsic variables:

- Air temperature
- Dew point temperature
- Wind direction
- Wind Speed

Derived Variables:

- Wet bulb temperature
- Past weather (lag of 5-, 10- and 15-min)

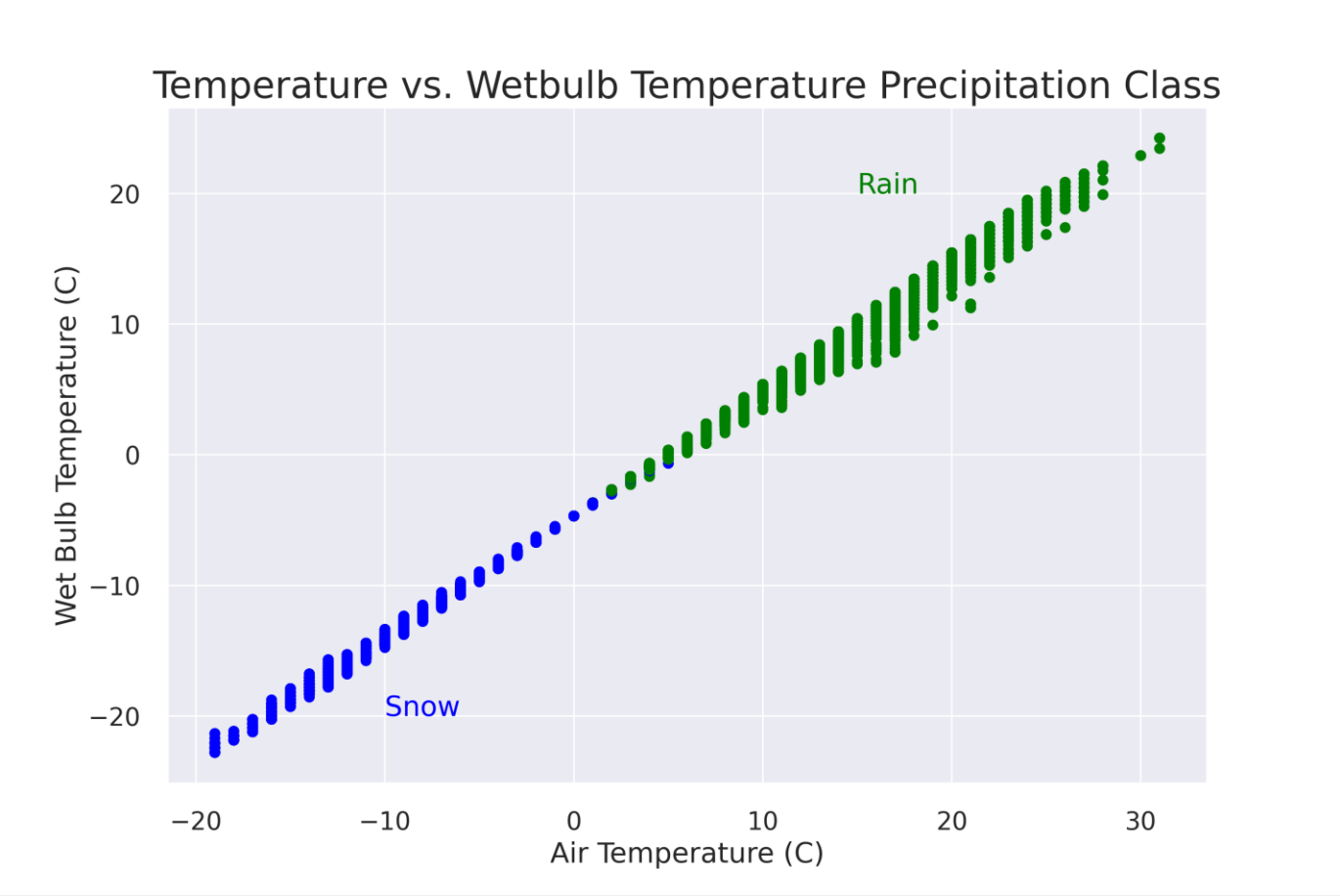


The Importance of Wet Bulb Temperature

T_w = Adiabatic saturation temperature

Between temperature and Dewpoint

Calculate using equation from Stull (2011)
- Temperature
- Relative Humidity (From Dewp)



SNOWathome.com

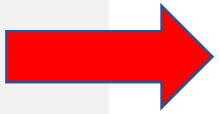
Wet-Bulb Temperature Chart

Fahrenheit

Good Snow Quality | Poor Snow Quality | No Snowmaking

Temp (F) \ Humidity	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%
20	14	14	14	15	15	15	16	16	16	17	17	18	18	18	19	19	19	20	20
21	14	15	15	16	16	16	17	17	17	18	18	18	19	19	19	20	20	21	21
22	15	16	16	16	17	17	17	18	18	19	19	19	20	20	20	21	21	22	22
23	16	16	17	17	18	18	18	19	19	19	20	20	21	21	21	22	22	22	23
24	17	17	18	18	18	19	19	20	20	20	21	21	22	22	22	23	23	23	24
25	18	18	18	19	19	20	20	20	21	21	22	22	22	23	23	24	24	24	25
26	18	19	19	20	20	20	21	21	22	22	23	23	23	24	24	25	25	25	26
27	19	19	20	20	21	21	22	22	23	23	23	24	24	25	25	26	26	26	27
28	20	20	21	21	22	22	23	23	23	24	24	25	25	26	26	27	27	27	28
29	20	21	21	22	22	23	23	24	24	25	25	26	26	27	27	28	28	28	29
30	21	22	22	23	23	24	24	25	25	26	26	27	27	28	28	29	29	29	30
31	22	22	23	23	24	25	25	26	26	27	27	28	28	29	29	29	30	30	31
32	23	23	24	24	25	25	26	26	27	27	28	28	29	29	30	30	31	31	32
33	23	24	24	25	26	26	27	27	28	28	29	29	30	30	31	31	32	32	33
34	24	25	25	26	26	27	27	28	29	29	30	30	31	31	32	32	33	33	34
35	25	25	26	27	27	28	28	29	29	30	31	31	32	32	33	33	34	34	35
36	25	26	27	27	28	29	29	30	30	31	31	32	33	33	34	34	35	35	36
37	26	27	27	28	29	29	30	31	31	32	32	33	34	34	35	35	36	37	37
38	27	27	28	29	29	30	31	31	32	33	33	34	35	35	36	36	37	38	38
39	27	28	29	30	30	31	32	32	33	34	34	35	35	36	37	37	38	39	39
40	28	29	30	30	31	32	32	33	34	34	35	36	36	37	38	38	39	40	40

Higher Temperature
Lower RH

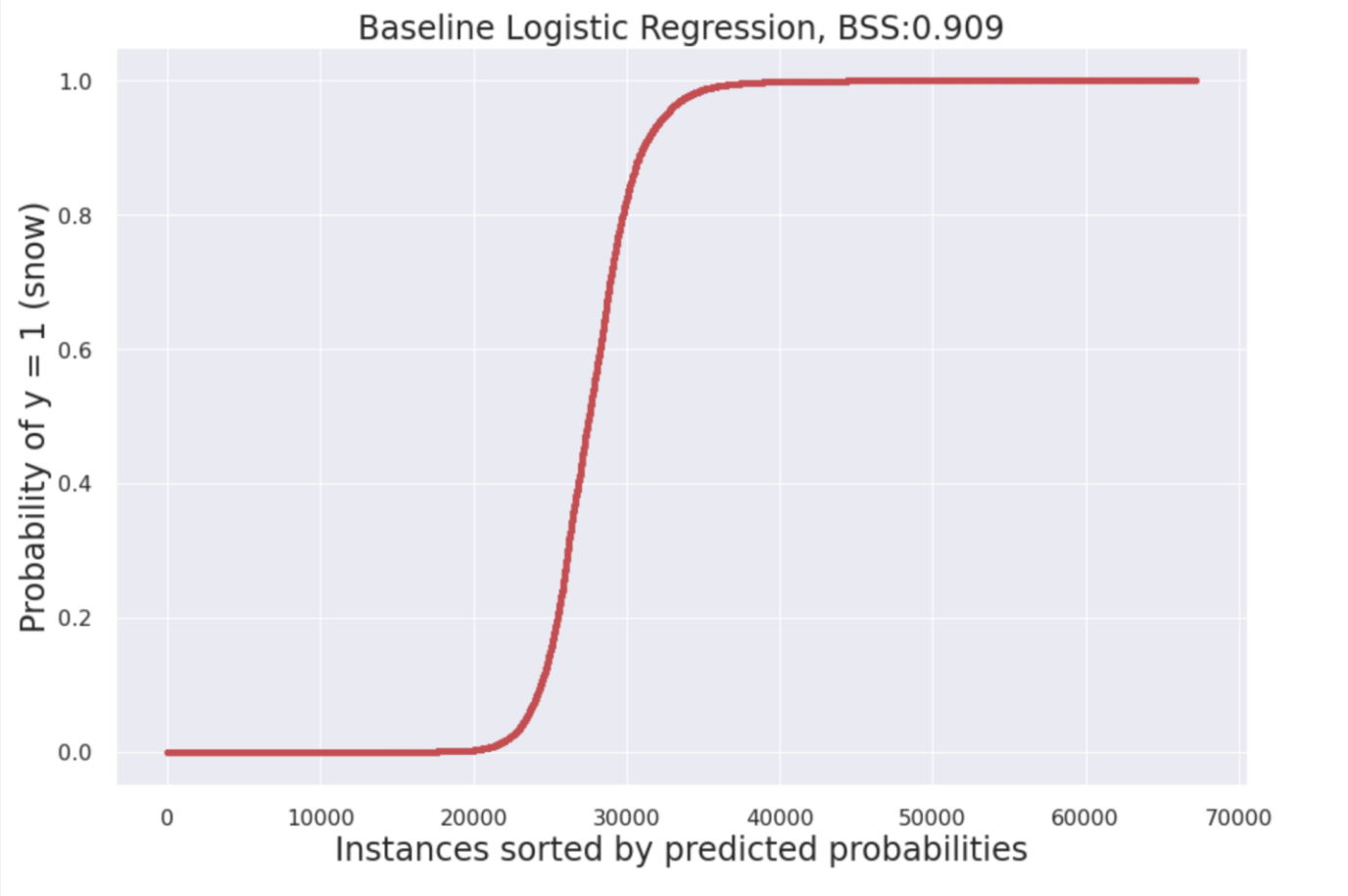


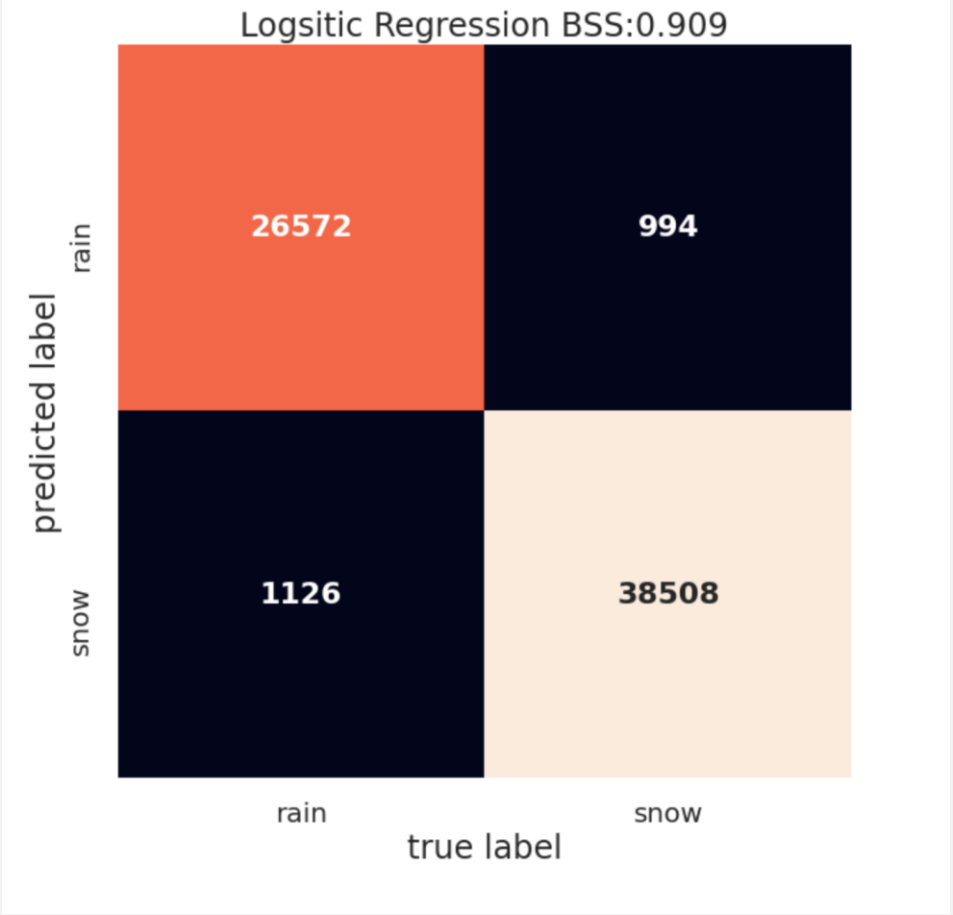
Wet bulb temperature is the lowest temperature that can be obtained by evaporating water into the air at a constant pressure. The term comes from the technique of wrapping a wet cloth around a mercury bulb thermometer and blowing air over the cloth until the water evaporates. The wet bulb temperature is always lower than the dry bulb temperature, but will be identical with 100% relative humidity.
This wet bulb temperature is what snowmakers use to know when they can make snow. You can see it is possible to make snow when the temperatures are above freezing but only with very low humidity.

Plot your current temperature (*red numbers on the left*) to the % of humidity *(blue numbers on the top)* and where they meet the *(black numbers)* is your current wet bulb temp.
Any time the wet bulb number is below 20 degrees Fahrenheit (blue shaded area) snowmaking is at its best… nice dry snow. You can make snow from 21 degrees to 27 degrees wet bulb *(purple shaded area)* but the snow will be wet.

Source: https://www.snowathome.com/pdf/wet_bulb_chart_fahrenheit.pdf

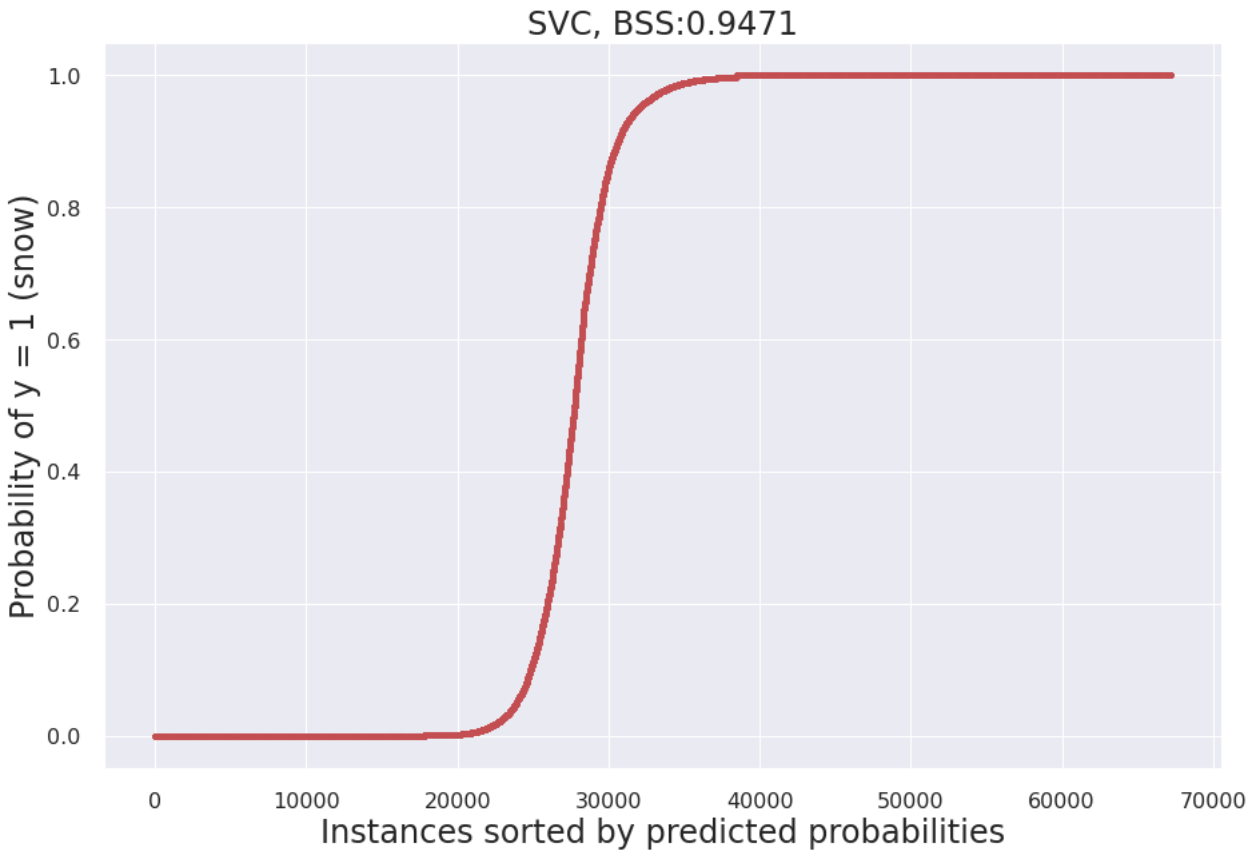
Logistic Regression (Baseline)

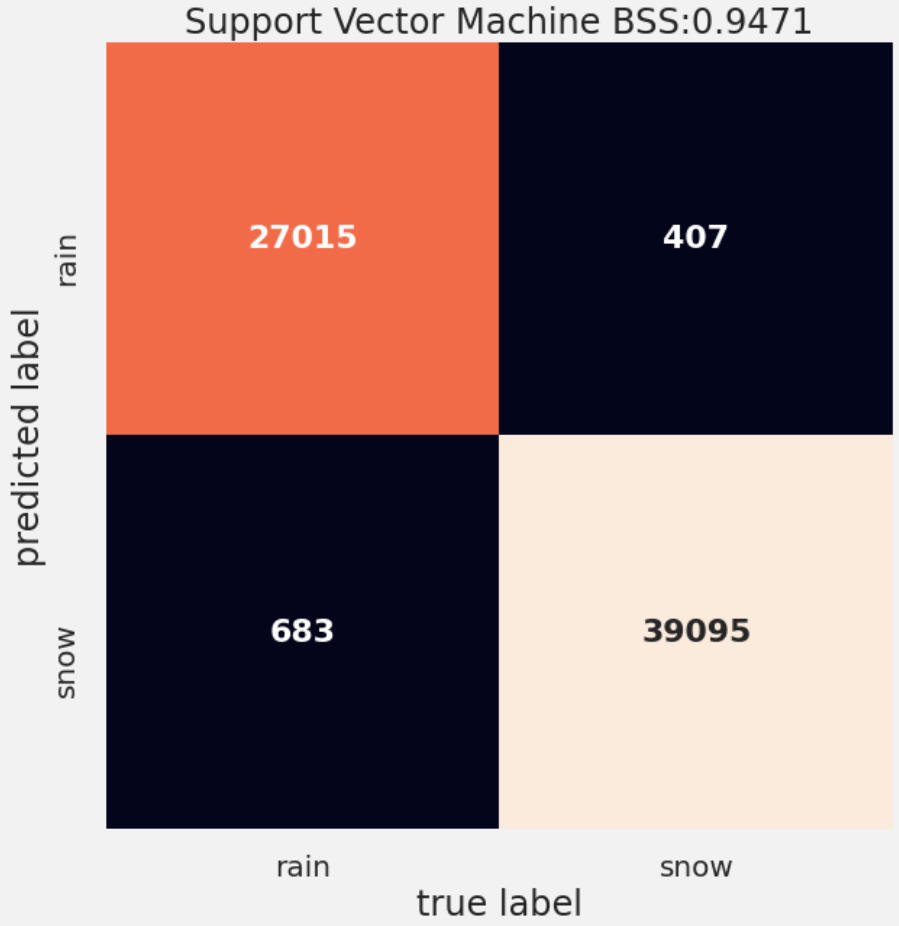




Support Vector Machine (SVM)

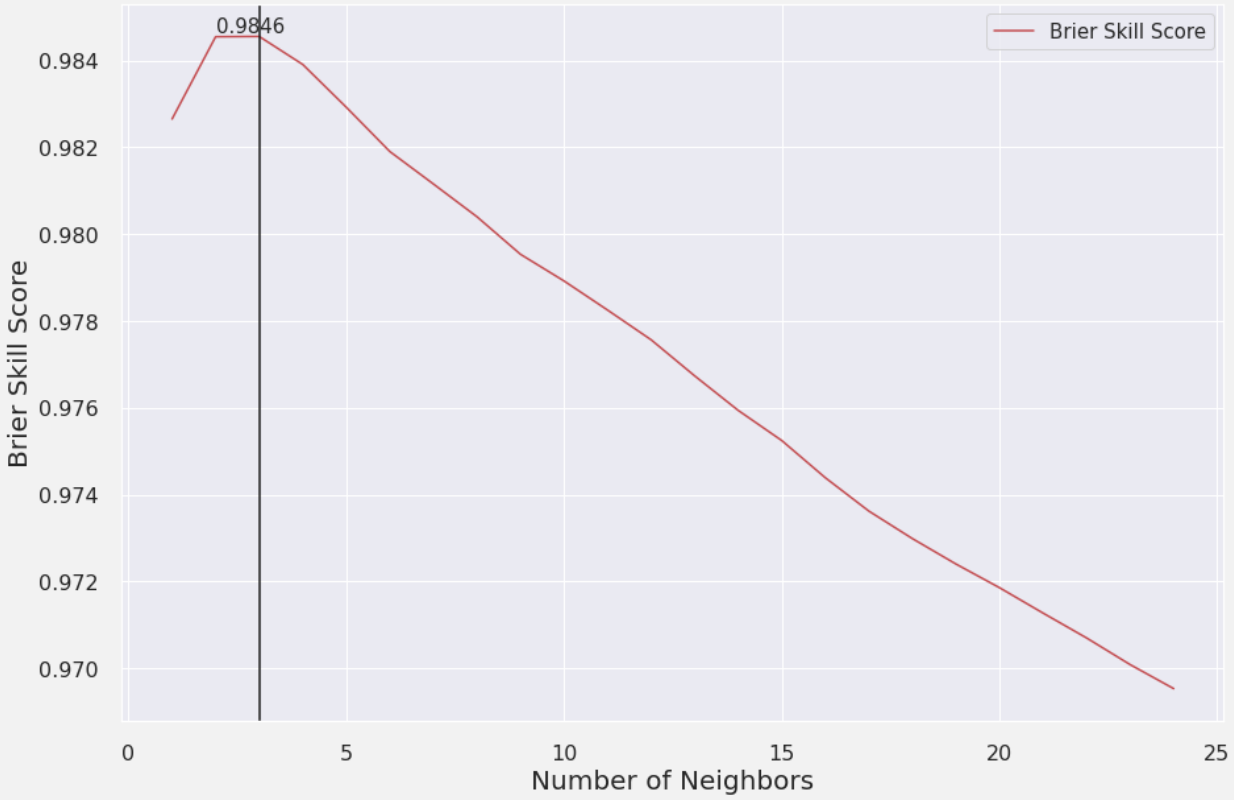
- Tried to optimize but took enormous computation time.

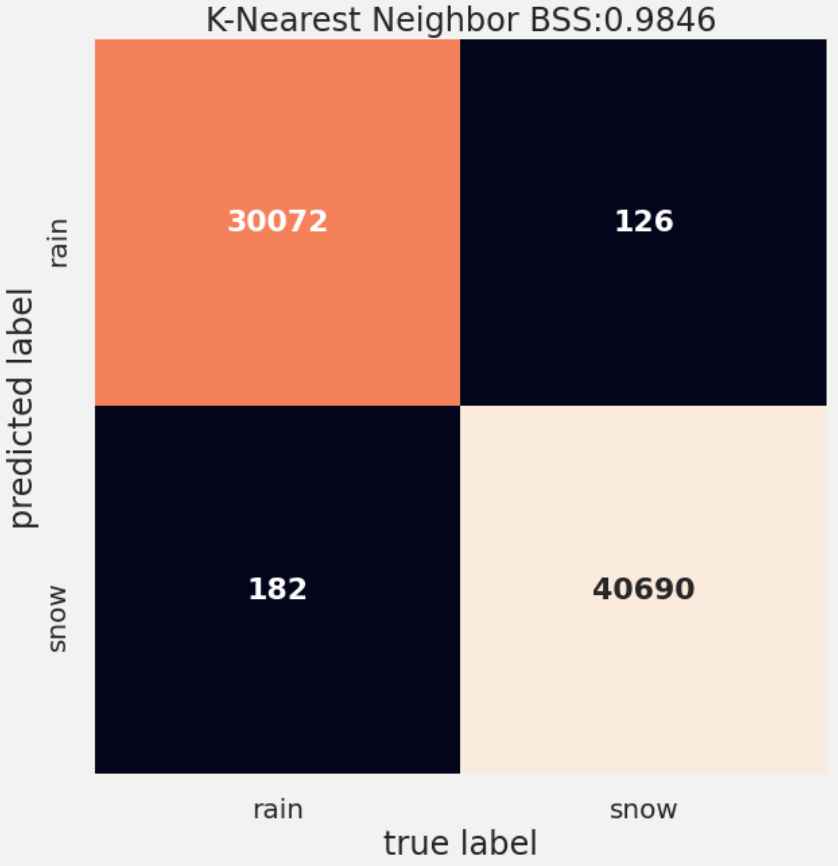


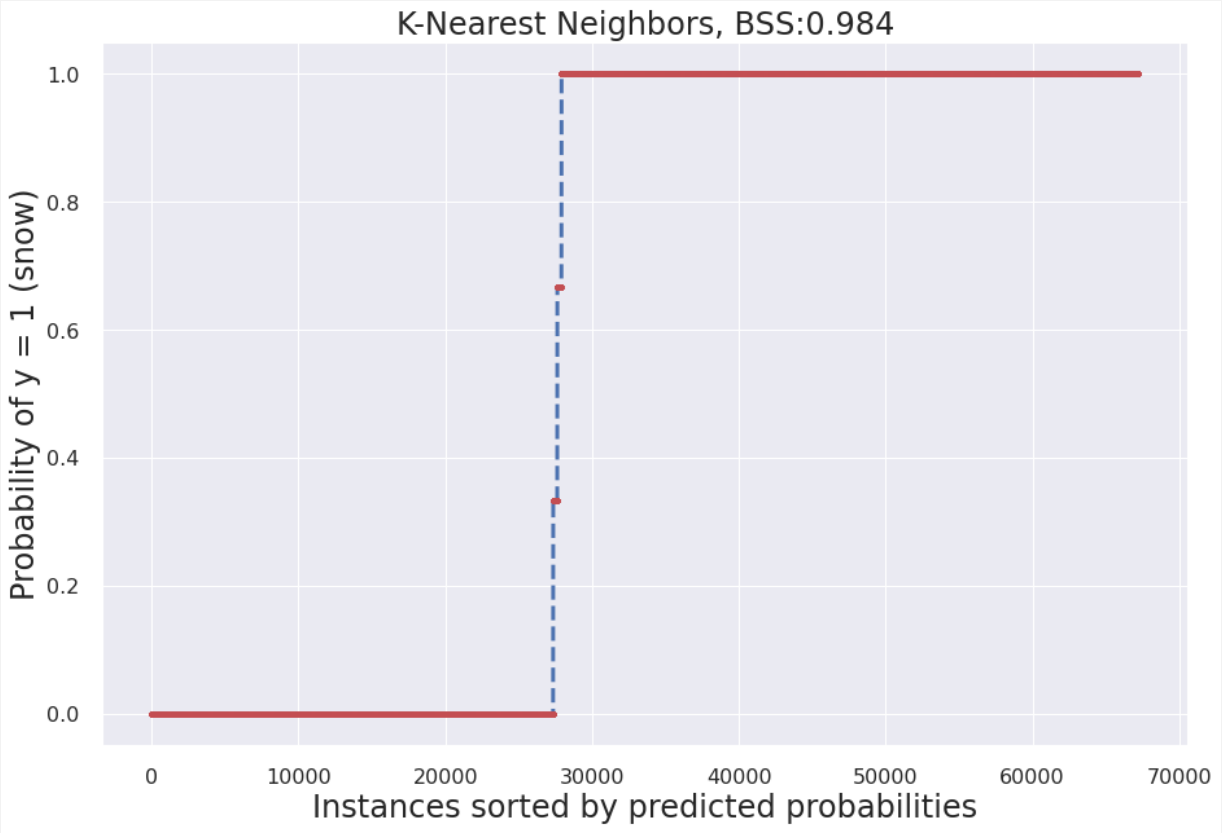


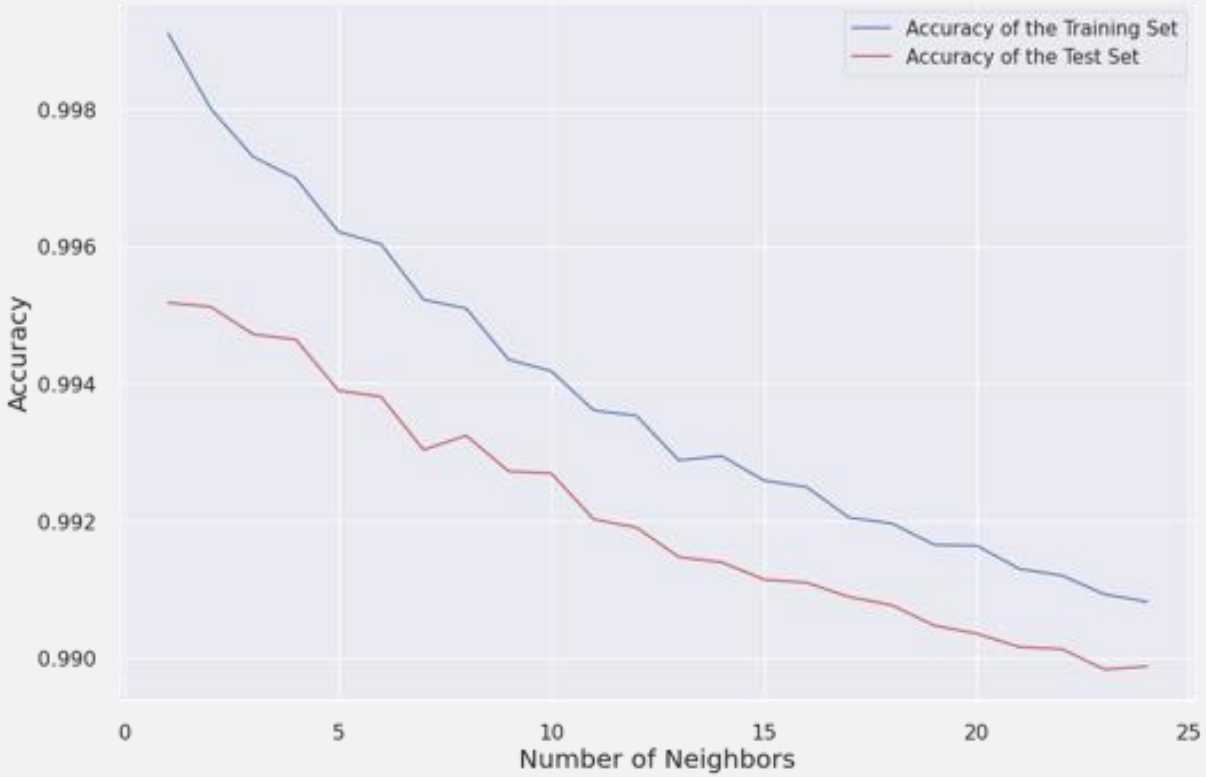
K-Nearest Neighbors (KNN)

- Optimized using n_neighbors hyperparameter






K-Nearest Neighbors, BSS:0.984
Probability of y = 1 (snow)
1.0
0.8
0.6
0.4
0.2
0.0
0
10000
20000
30000
40000
50000
60000
70000
Instances sorted by predicted probabilities


Accuracy of the Training Set
Accuracy of the Test Set
Accuracy
0.998
0.996
0.994
0.992
0.990
0
5
10
15
20
25
Number of Neighbors

Gaussian Naïve Bayes (NB): Base

Methods used

- Original NB
- Calibration using isotonic and sigmoid methods

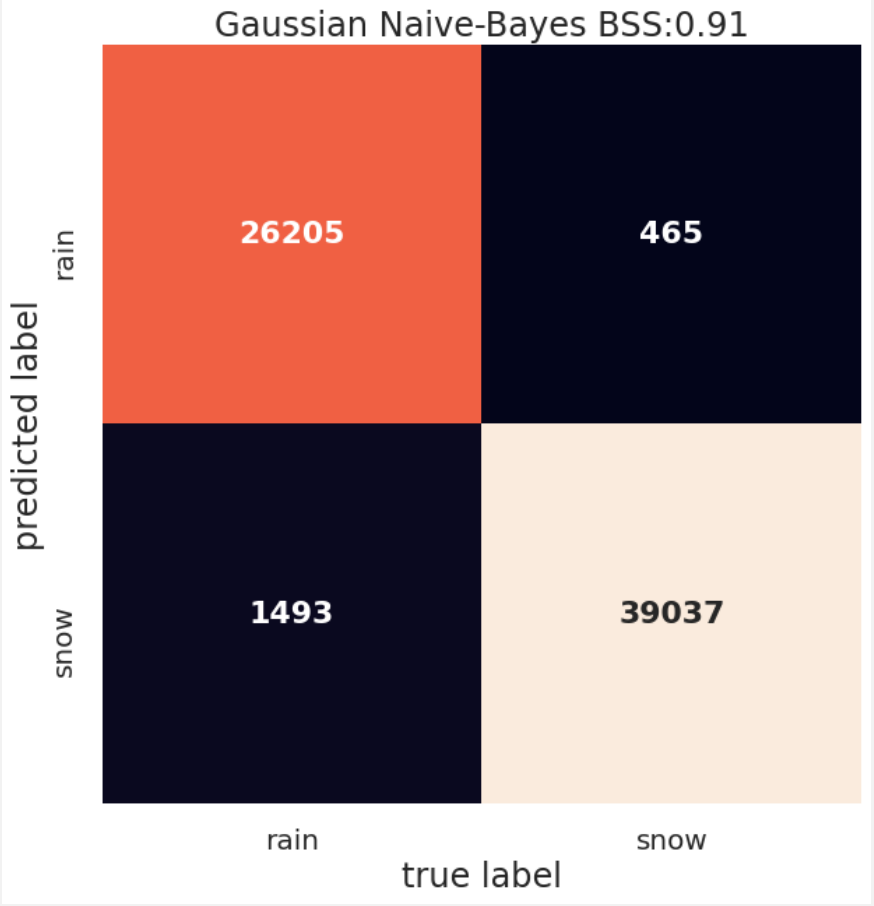
```
from sklearn.calibration import CalibratedClassifierCV

clf_isotonic = CalibratedClassifierCV(model, cv=2, method='isotonic')
```

Variables

Starting with 4 intrinsic variables + wet bulb

→ BSS = 0.91

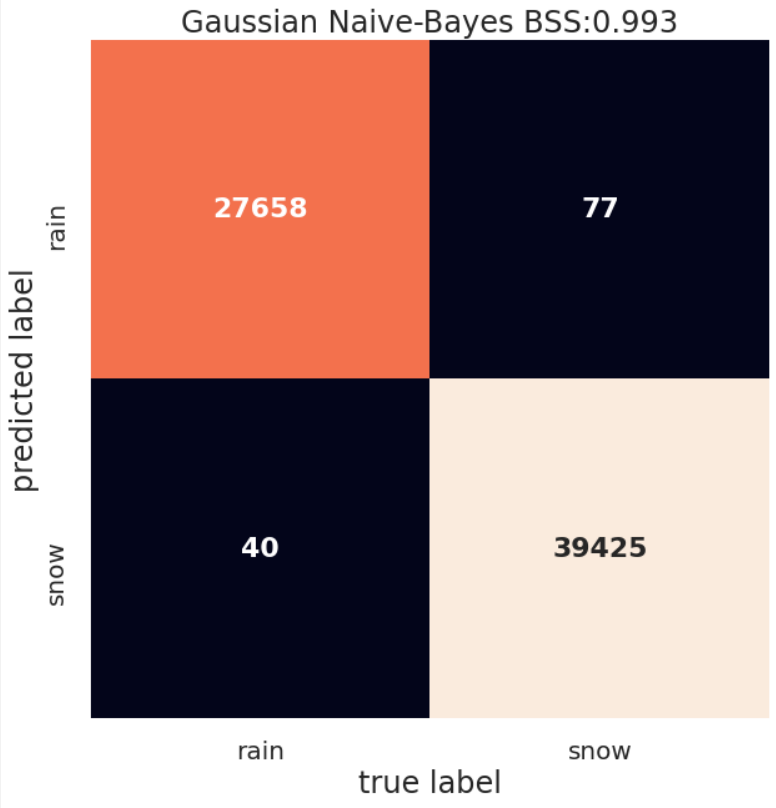


Confusion matrix for calibrated Gaussian Naïve Bayes using 4 intrinsic variables + derived wet bulb temperature

Gaussian NB: More Predictors

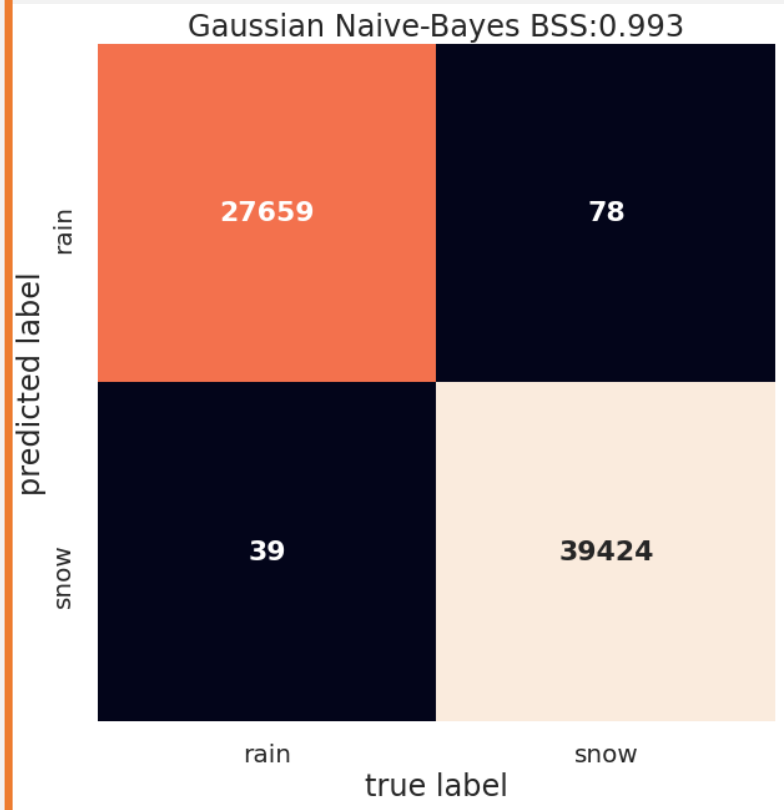
Vars: + lag5min

BSS: + 0.083 (9%)



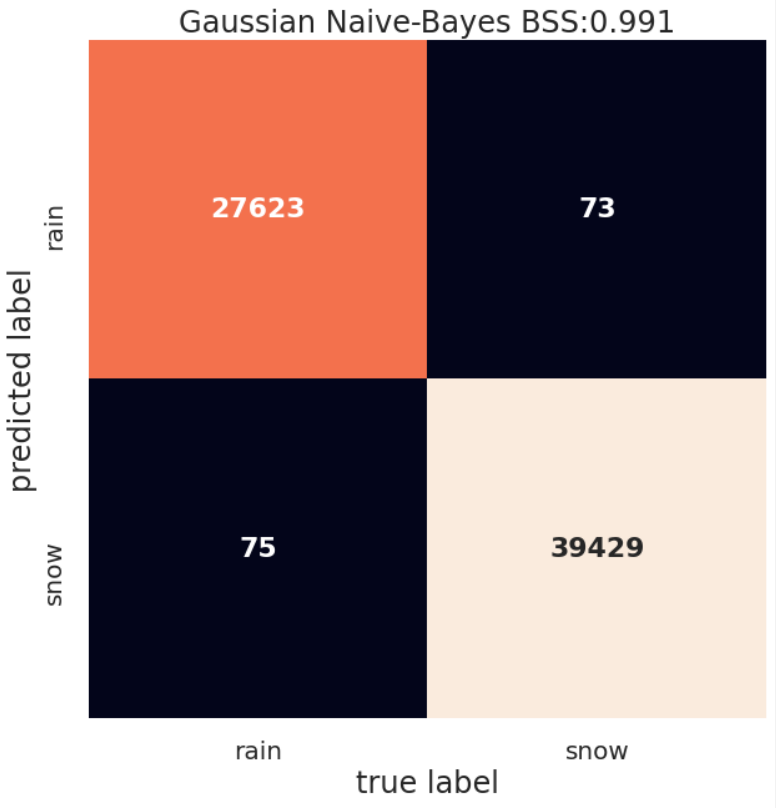
++ lag10min

+ 0.083 (9%)

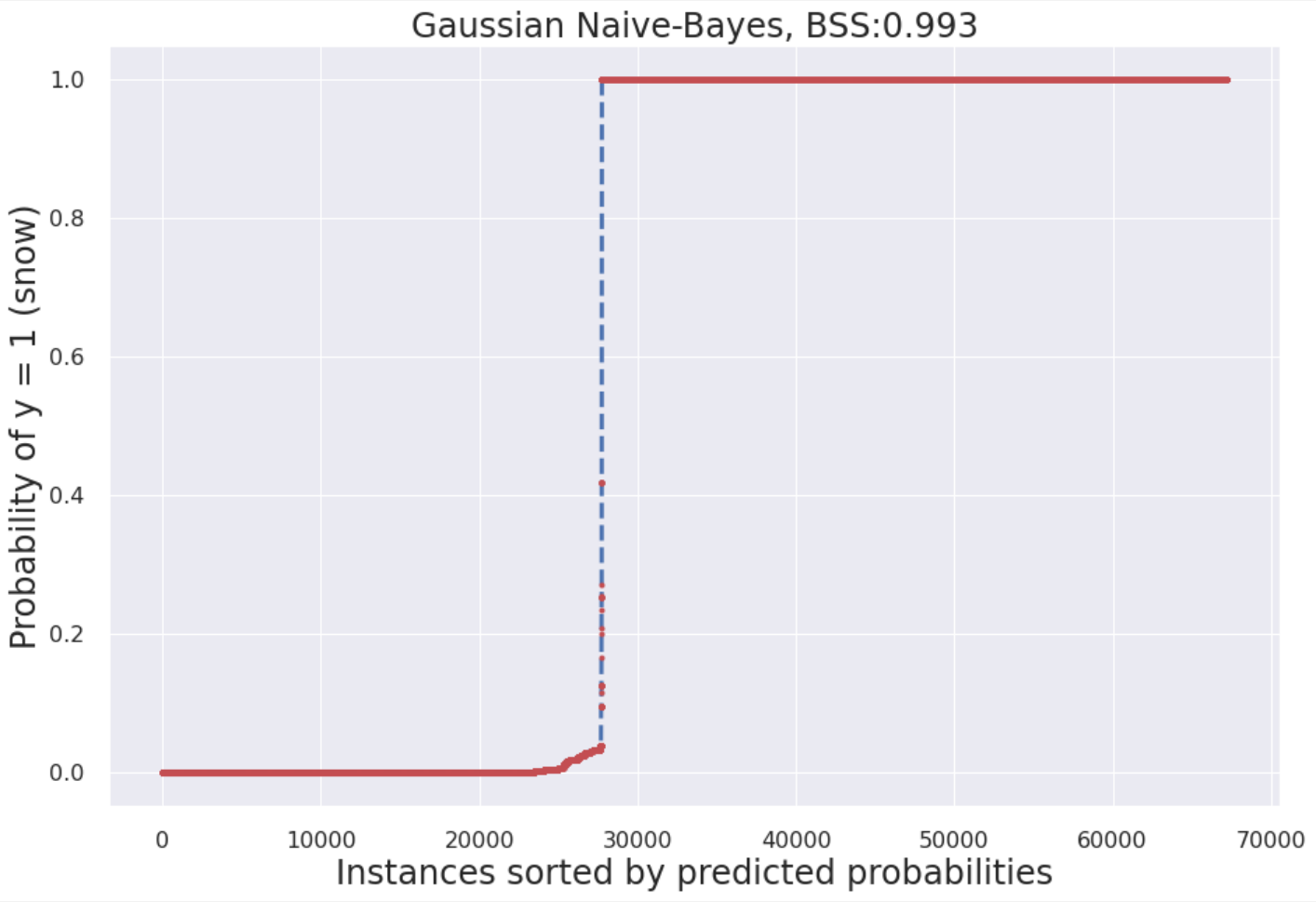


+++ lag15min

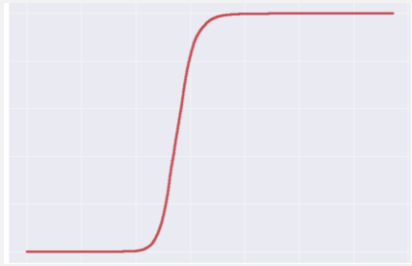
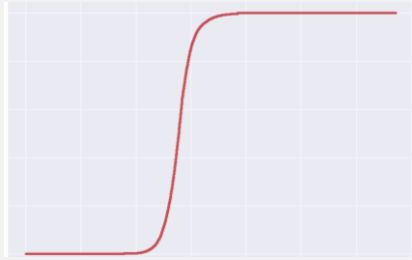
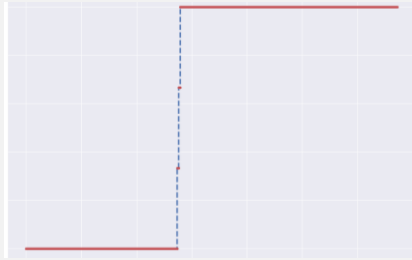
+ 0.081 (9%)



Gaussian NB: Probability distribution



Summary

Model 1	Logistic regression	BSS = 0.909	
	SVM (not optimized)	BSS = 0.947 (+4%)	
	KNN	BSS = 0.984 (+8%)	
Model 2	Gaussian NB	BSS = 0.993 (+9%)