

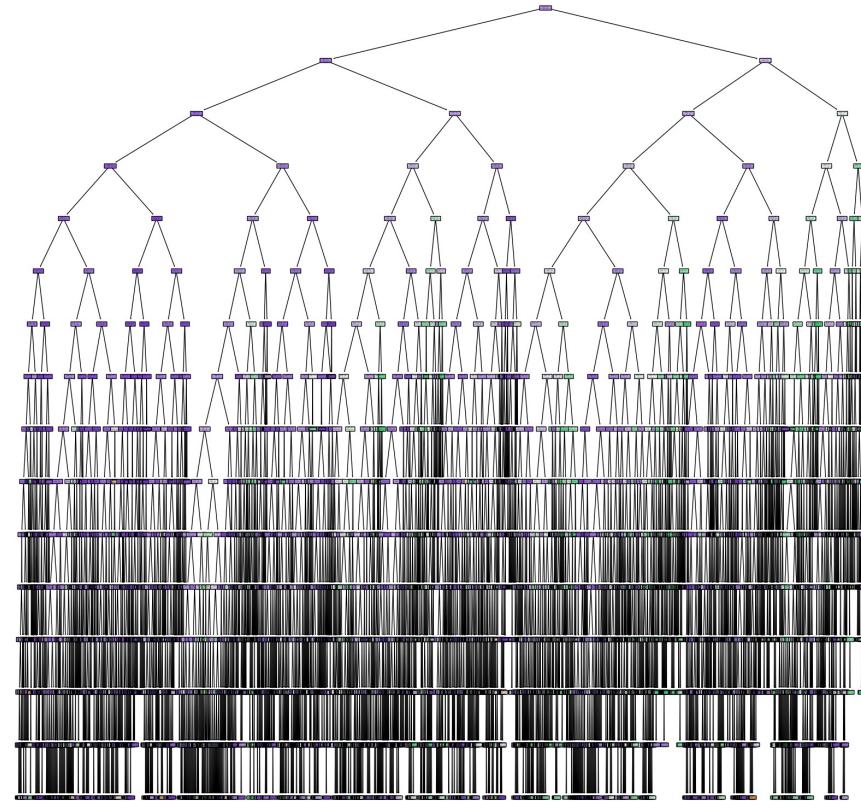
# DECISION TREE

---

Hosna Habibi

# AGENDA

- Data
- Entropy
- Gini
- Log\_Loss
- Pruning





	1	2	3	4	5	6	7	8	9	10	...	34	35	36	37	38	39	40	41	42	win
0	b	b	b	b	b	b	b	b	b	b	...	b	b	b	b	b	b	b	b	b	win
1	b	b	b	b	b	b	b	b	b	b	...	b	b	b	b	b	b	b	b	b	win
2	b	b	b	b	b	b	o	b	b	b	...	b	b	b	b	b	b	b	b	b	win
3	b	b	b	b	b	b	b	b	b	b	...	b	b	b	b	b	b	b	b	b	win
4	o	b	b	b	b	b	b	b	b	b	...	b	b	b	b	b	b	b	b	b	win

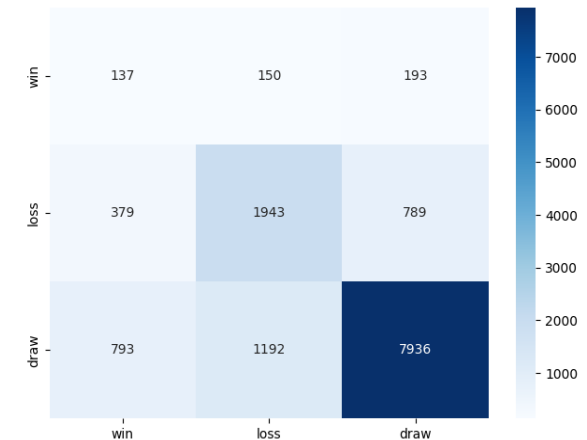
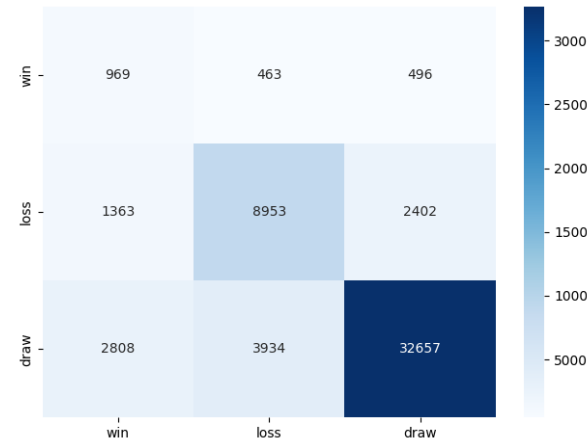
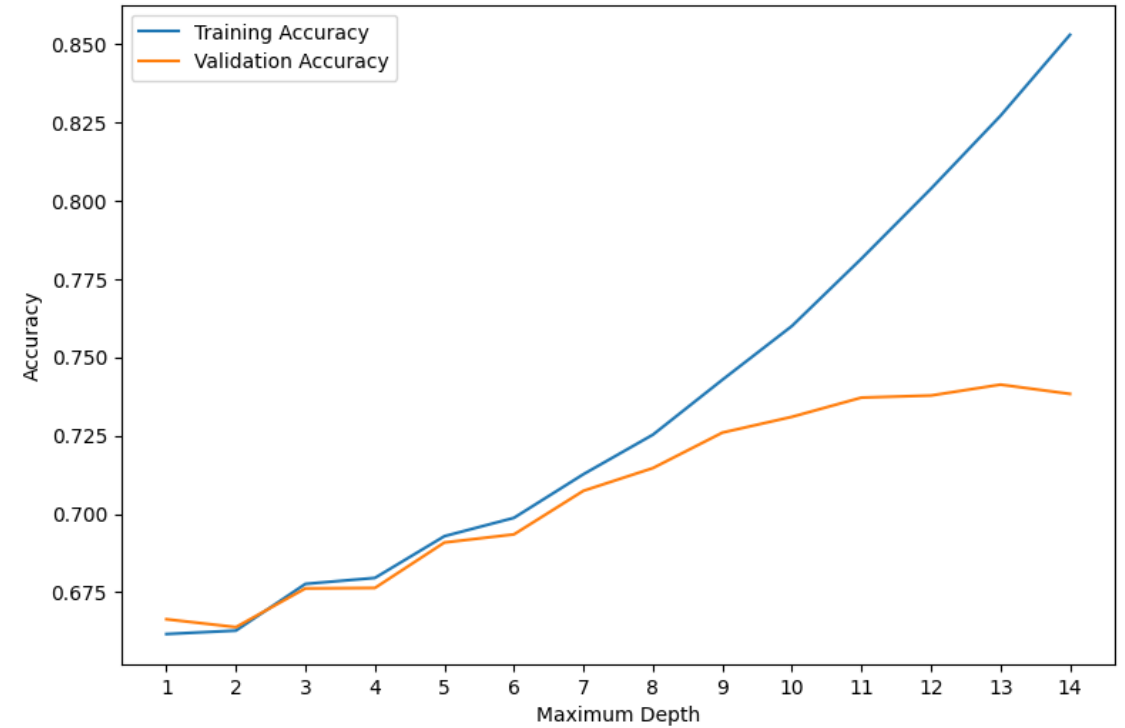
5 rows x 23 columns

# DATA

- Target Classes (Win, Draw, Loss)
- 42 features
- 67557 samples

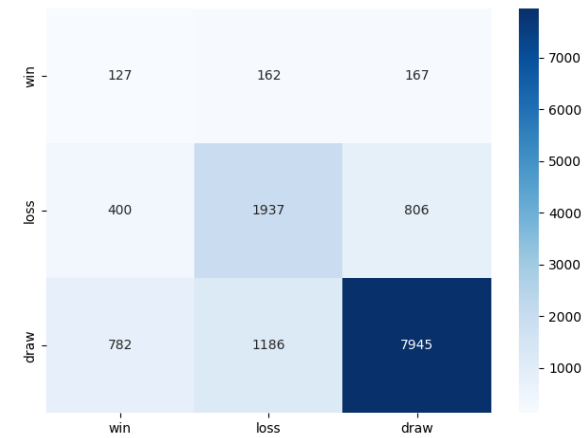
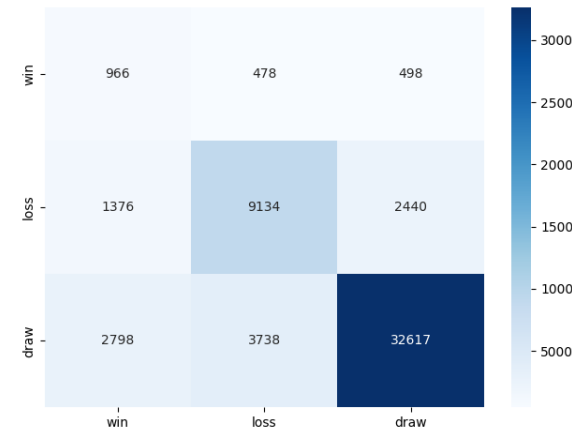
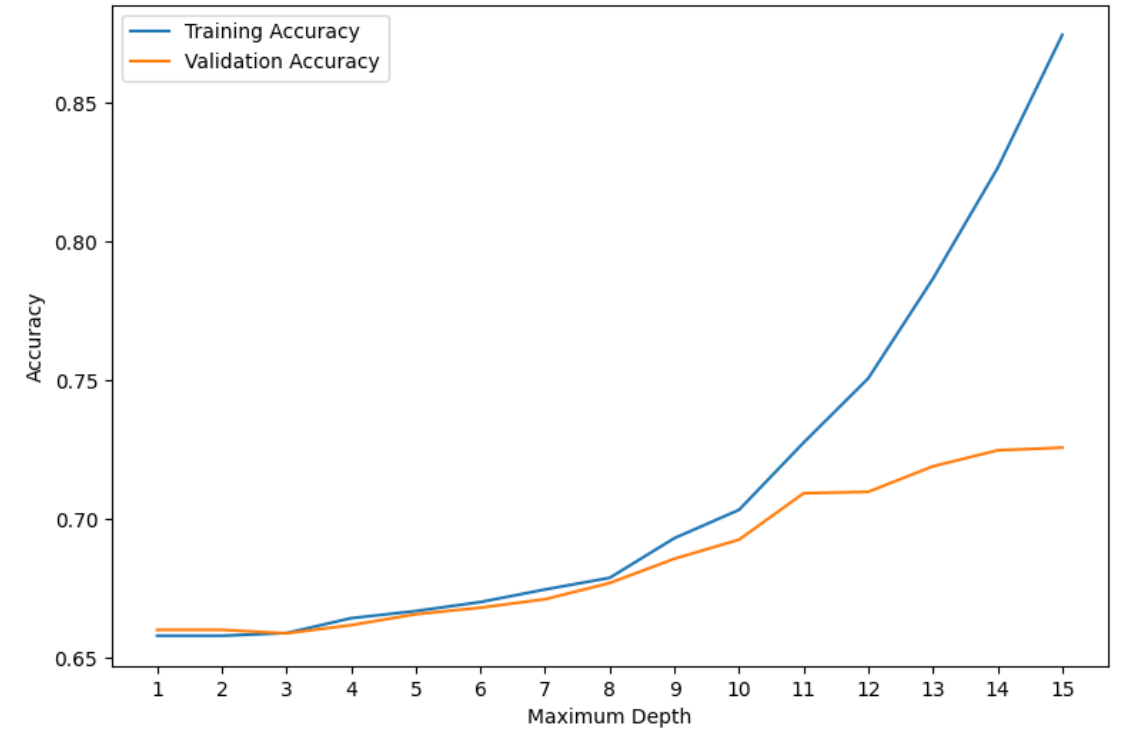
# ENTROPY

- 20-80 Splitting:
- 74% Test
- 78% Train
- 50-50 Splitting:
- 73% Test



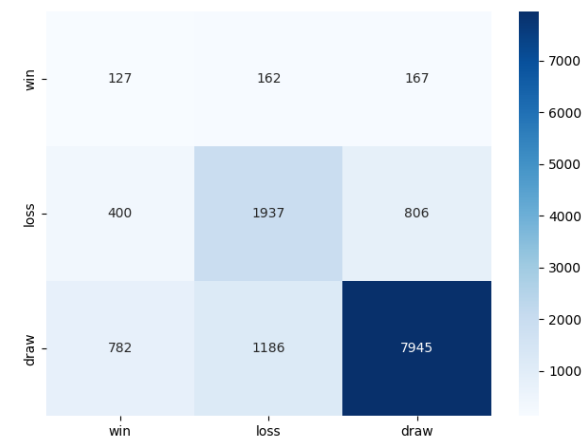
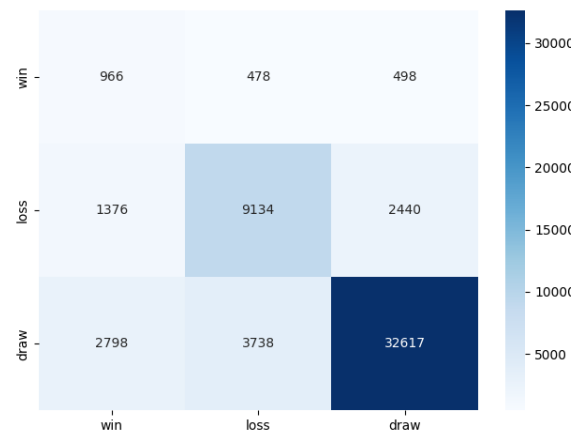
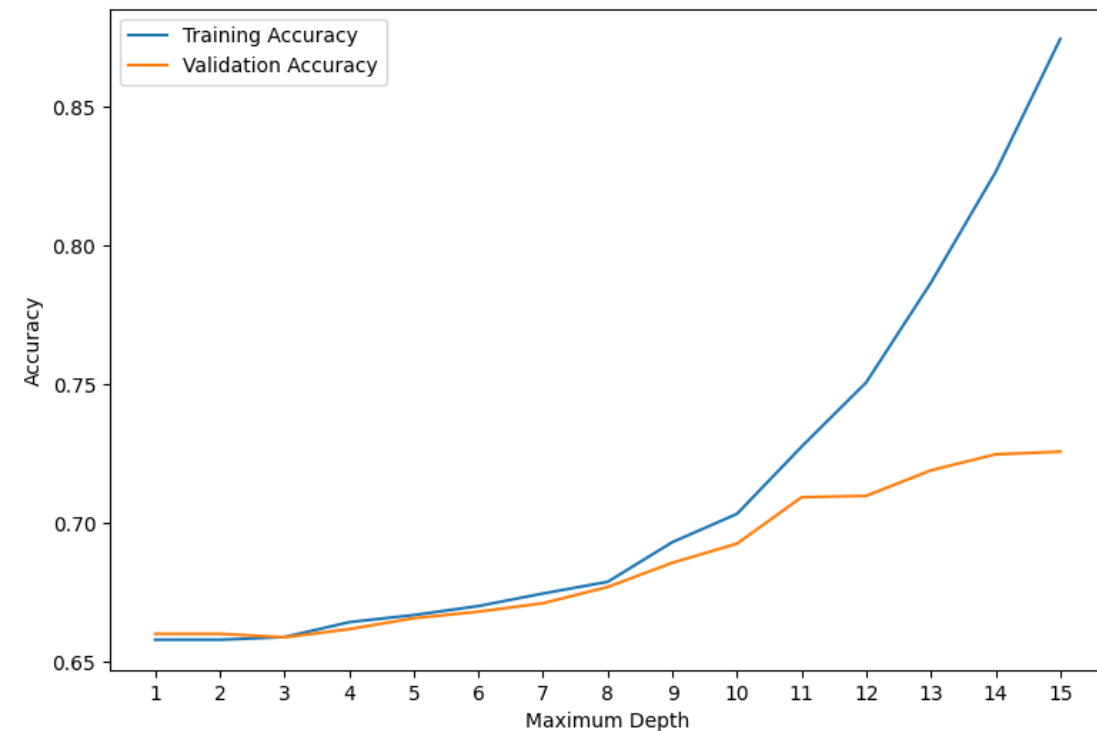
# GINI

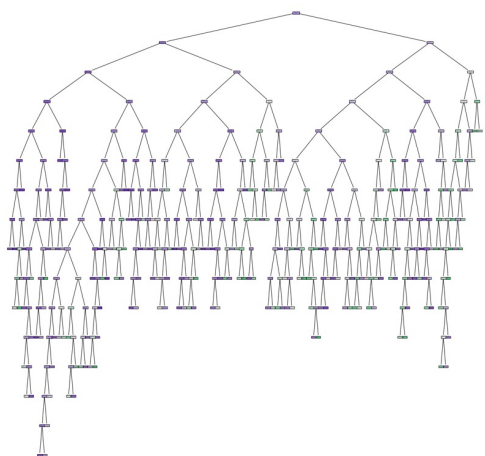
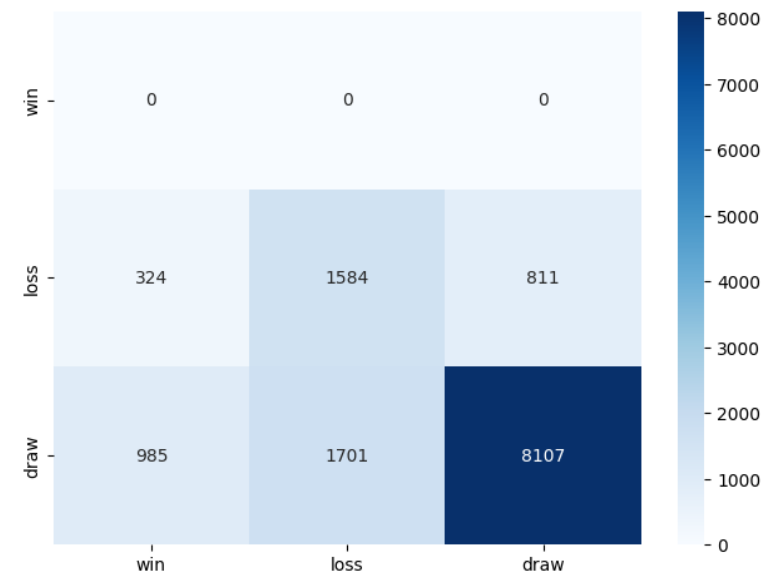
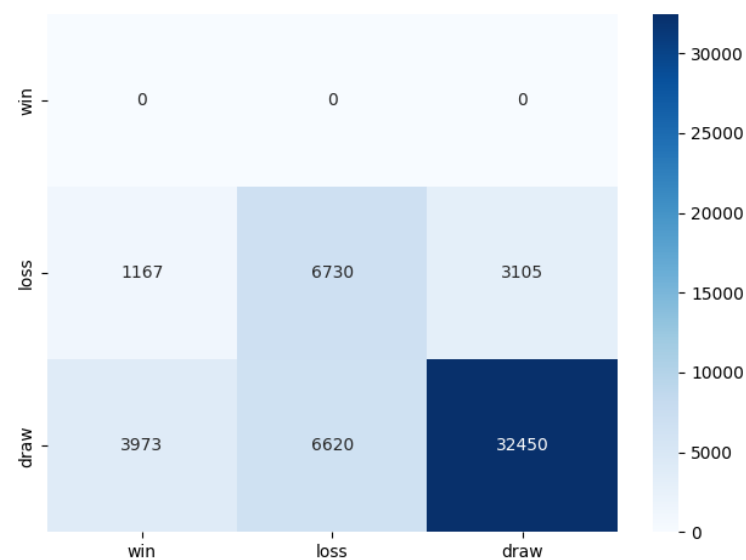
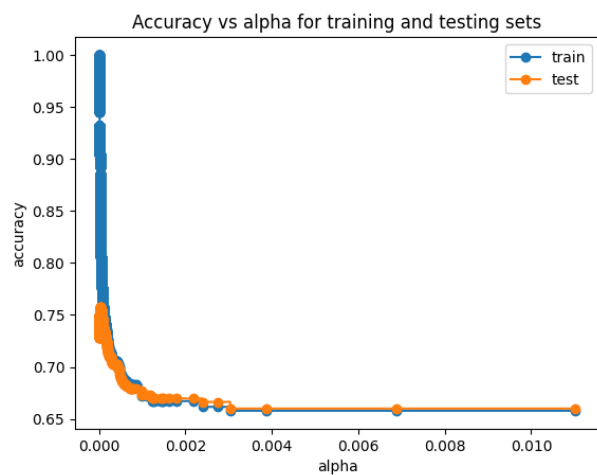
- 20-80 Splitting:
- 74% Test
- 79% Train
- 50-50 Splitting:
- 71% Test



# LOG\_LOSS

- 20-80 Splitting:
- 74% Test
- 79% Train
- 50-50 Splitting:
- 71% Test





# PRUNNING & GCV