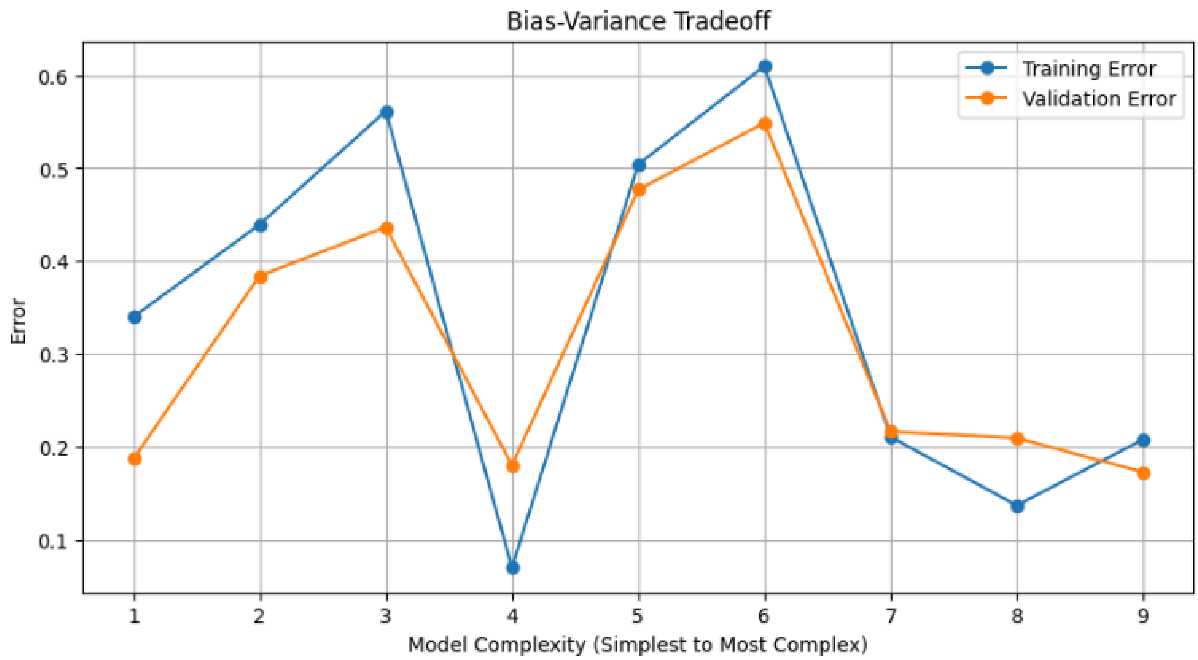
Models Validation Error

Model 1	0.1871
Model 2	0.3842
Model 3	0.4367
Model 4	0.1804
Model 5	0.4769
Model 6	0.5488
Model 7	0.2166
Model 8	0.2097
Model 9	0.1731

Final Winning Model

The final winning model, based on the validation dataset, is Model 7, which achieved low validation error among all models, with a pAUC-aboveTPR of 0.1034 and a recall rate of 0.64 for class 1. Model 7, an EfficientNet, likely outperformed simpler models like the 3-layer CNN and ResNet due to its greater complexity. However, I realized that complex architectures like EfficientNet and ResNet often require data preprocessing and normalization tailored specifically to their structure. By applying the same preprocessing and normalization for Model 1 to 3 to all models, I may have limited the performance of the complex models, which could explain why Model 1, the simpler model, performed almost as well as Model 7.

Bias-variance tradeoff



The chart suggests that models at complexity levels 4 and 7-9 offer the best balance between bias and variance. However, to fully leverage more complex models, specific tuning or adjustments might be necessary. Complexity levels 5 and 6 likely introduce excessive variance or training instability, making them unsuitable for this task.

Winning Model Performance

```
The partial auroc of the final model on the test image is 0.13625176183538829
              precision    recall  f1-score   support

     Class 0       0.98      0.92      0.95      1431
     Class 1       0.24      0.61      0.35        59

    accuracy                           0.91      1490
   macro avg       0.61      0.77      0.65      1490
weighted avg       0.95      0.91      0.93      1490
```

Test Data Performance

Based on the test dataset metrics, the model achieved a recall rate of 0.61 for Class 1 and a pAUC-aboveTPR of 0.136, both of which are higher than the validation set metrics. Given the constraints in environment and computing power, these results are satisfactory. However, performance could likely be improved with tailored data preprocessing and normalization specifically suited for EfficientNet. Despite these positive results, I believe the model still isn't optimal for this prediction task, as achieving higher performance would likely require more data for Class 1. This would help the model better capture the patterns necessary to increase recall and overall accuracy in the target class.

Validation, and test performance metrics of the winning model

	validation	test
recall	0.64	0.61
pAUC-aboveTPR	0.1034	0.136

The recall rate decreased from 0.64 on the validation set to 0.61 on the test set, while the pAUC-aboveTPR increased from 0.1034 on the validation set to 0.136 on the test set.