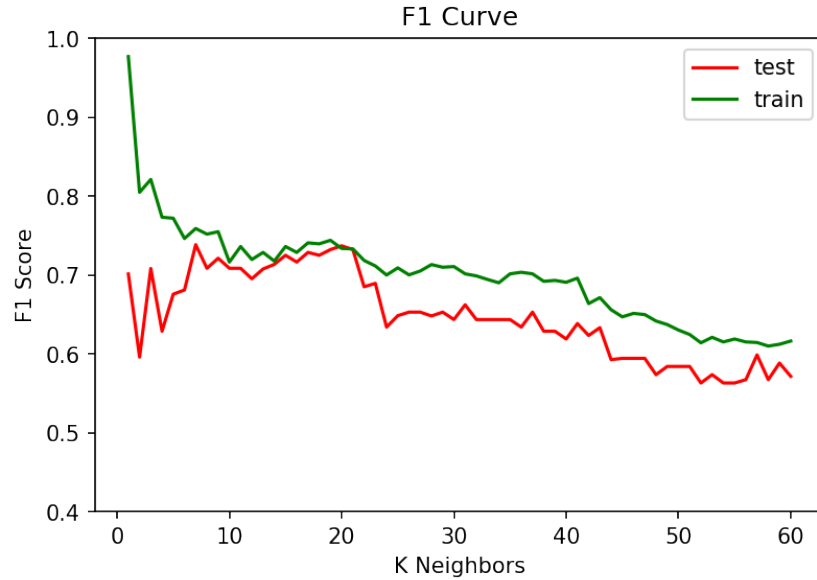
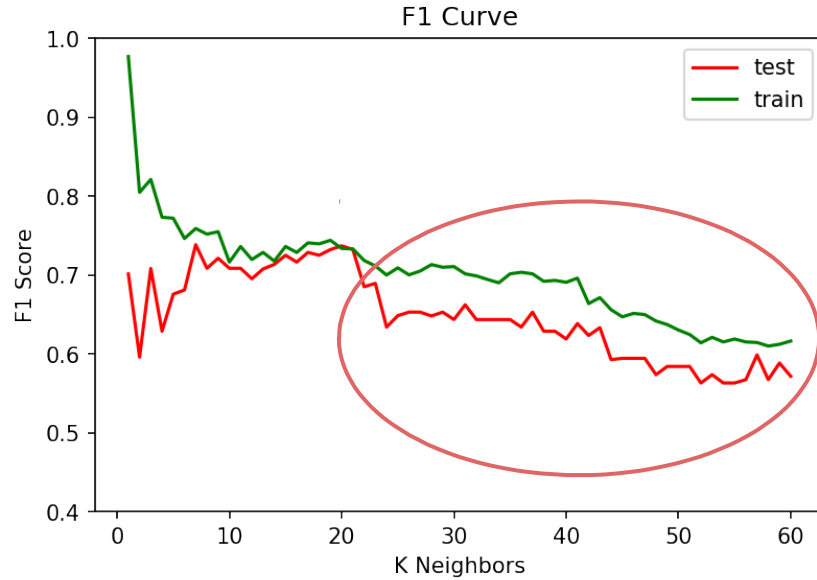


Bias and Variance in Predictive Modelling

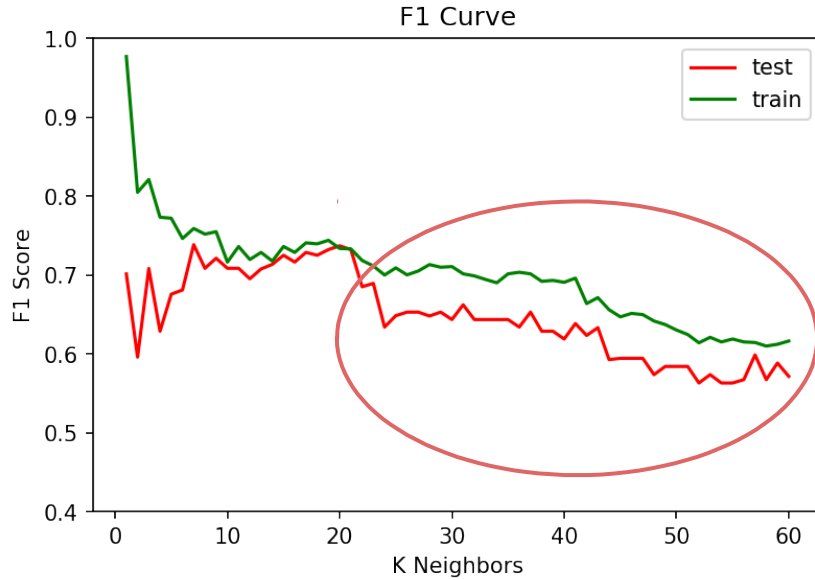
Bias and Variance



Bias and Variance



Bias and Variance

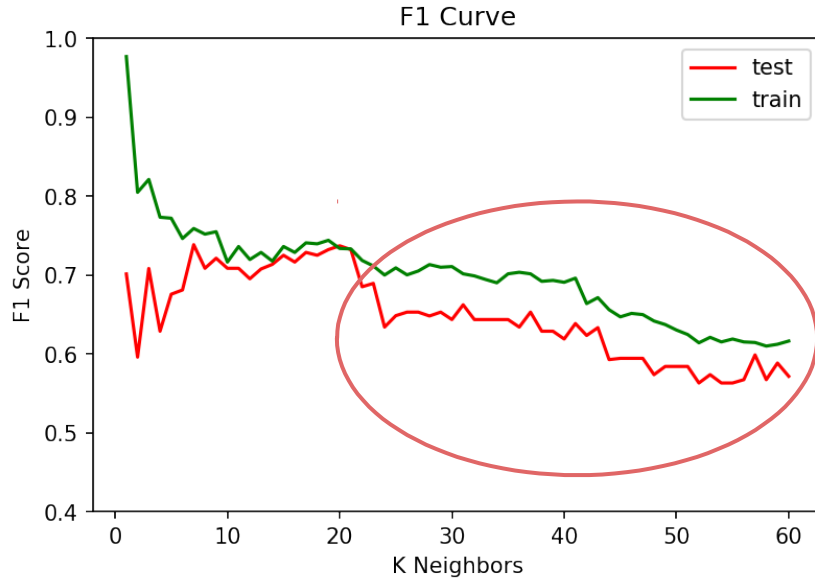


A

Not interested in learning

Class test ~50%
Test ~47%
Under-fitting

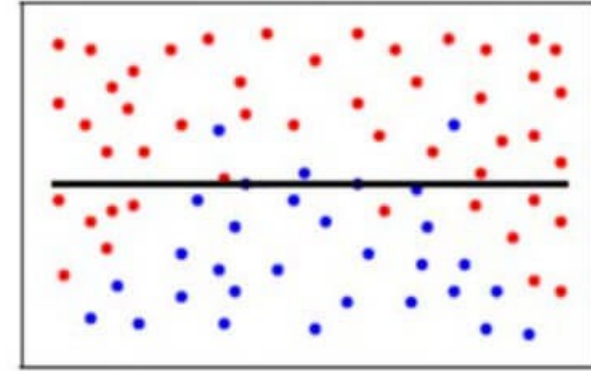
Bias and Variance



A

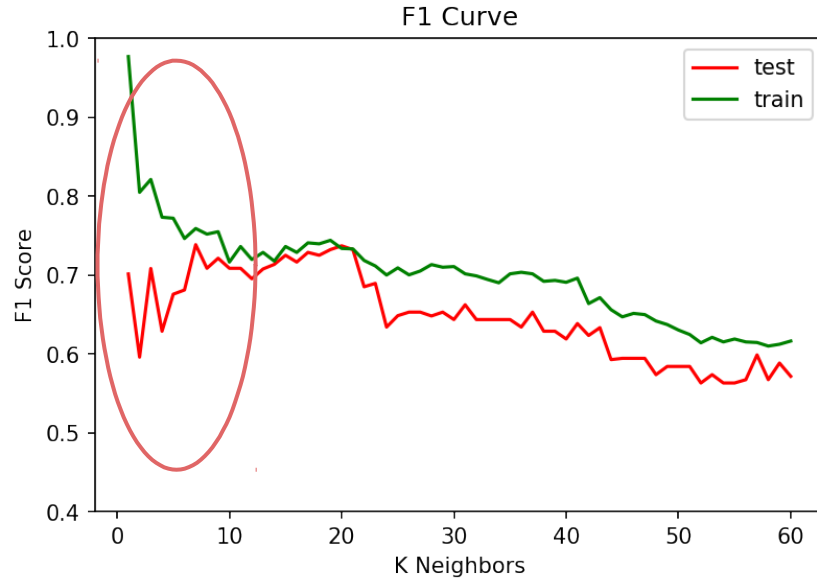
Not interested in learning

Class test ~50%
Test ~47%
Under-fitting

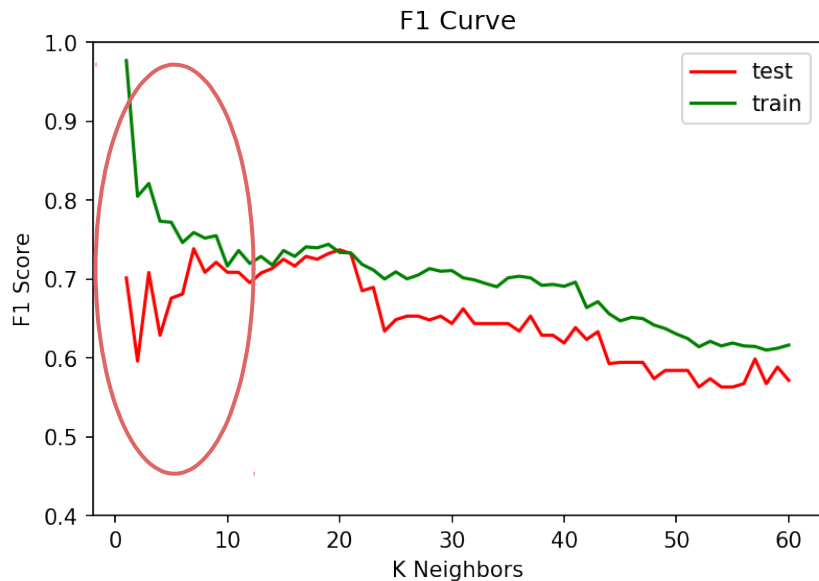


Training Set

Bias and Variance



Bias and Variance



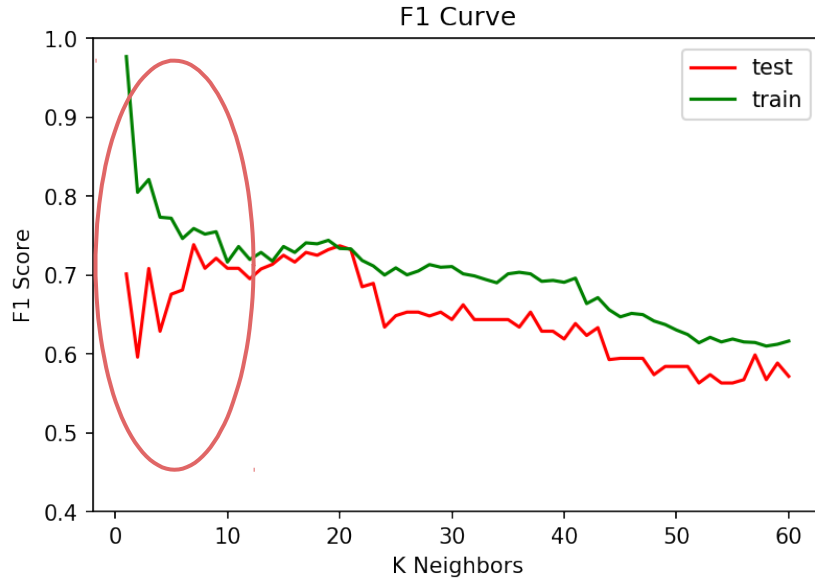
B

Memorising the lessons

Class test ~98%
Test ~52%

Over-fitting

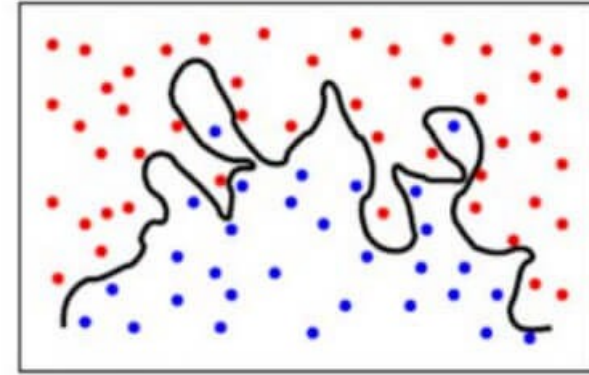
Bias and Variance



Memorising the lessons

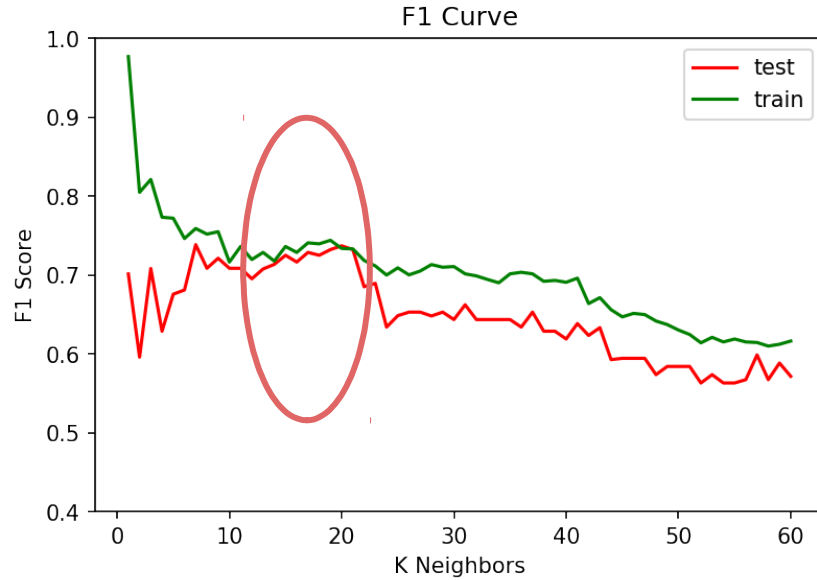
Class test ~98%
Test ~52%

Over-fitting

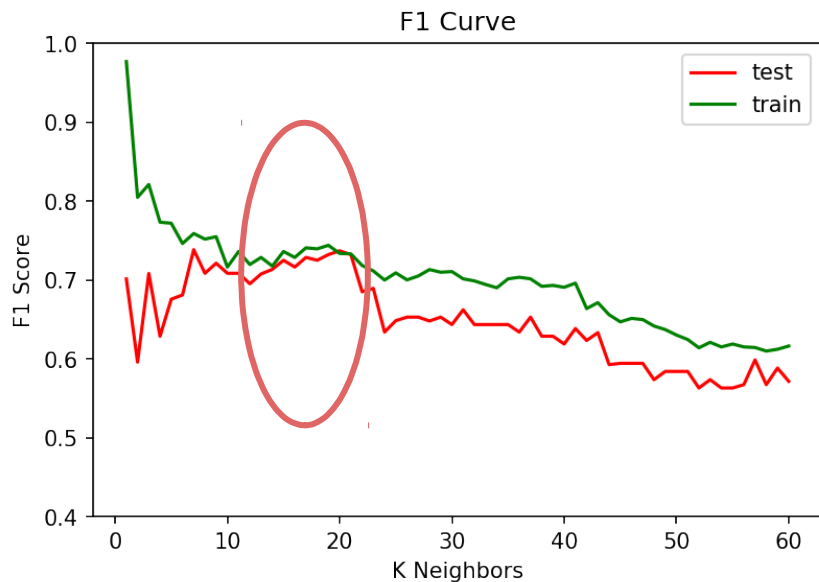


Training Set

Bias and Variance



Bias and Variance

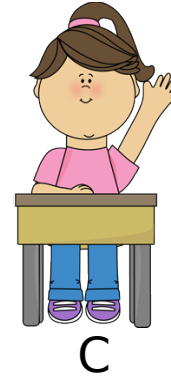
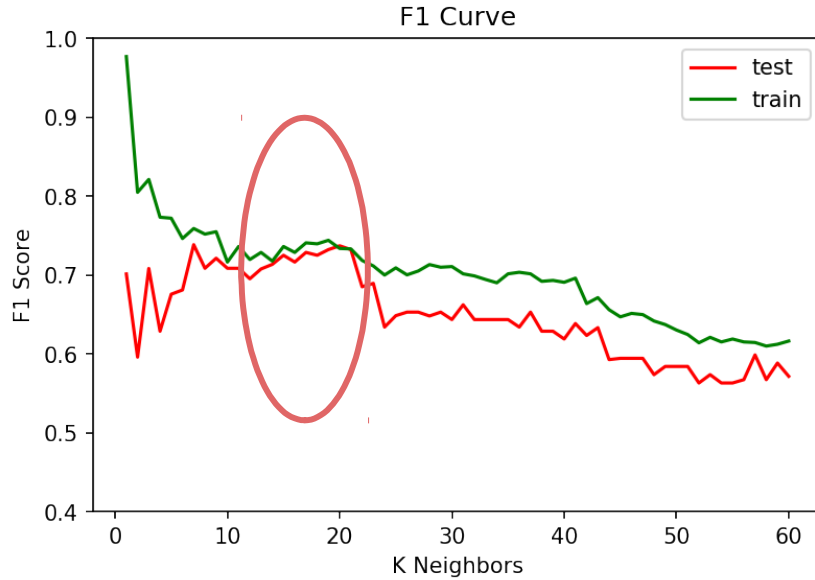


Consistent
Performance

Class test ~87%
Test ~84%

Best Fit

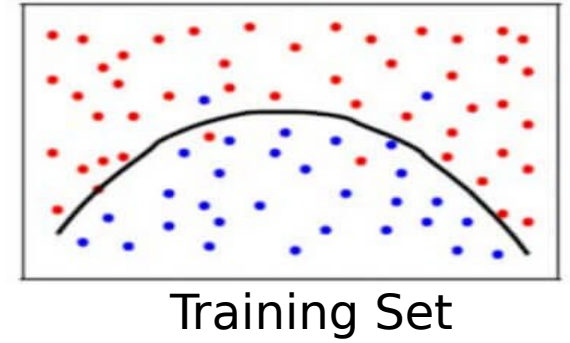
Bias and Variance



Consistent
Performance

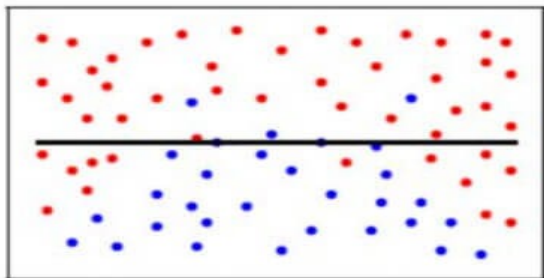
Class test ~87%
Test ~84%

Best Fit

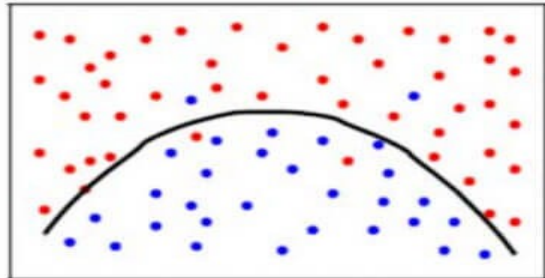


Bias and Variance

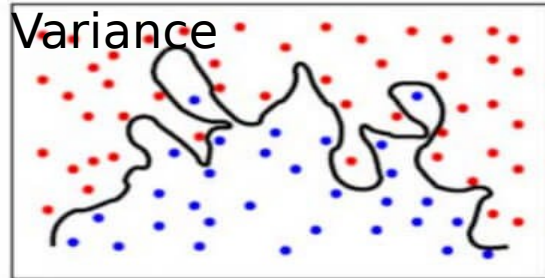
High Bias



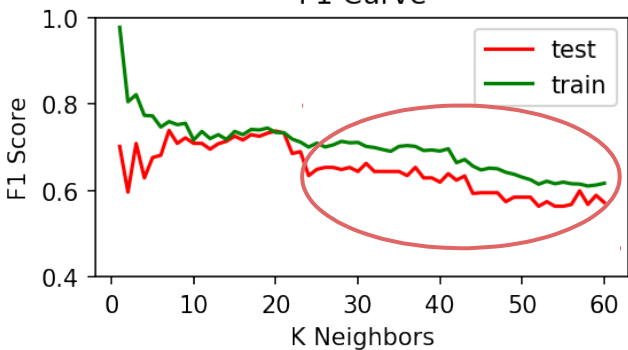
Optimum



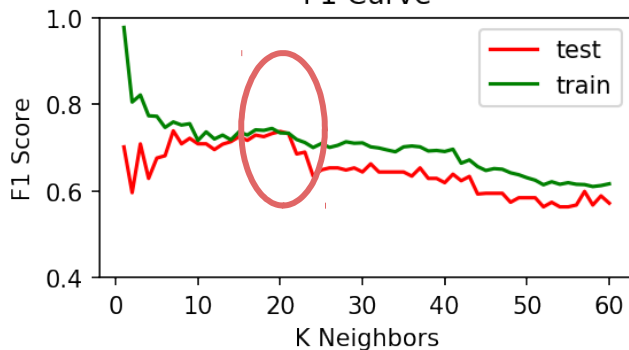
High Variance



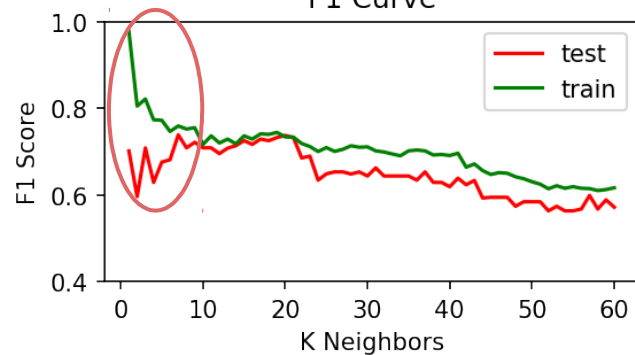
F1 Curve



F1 Curve



F1 Curve



Bias and Variance Summary:



A

Not interested in learning

Class test ~50%
Test ~47%

Under-fit/ less learning



B

Memorizing the lessons

Class test ~98%
Test ~69%

Over-fit/ Memorizing

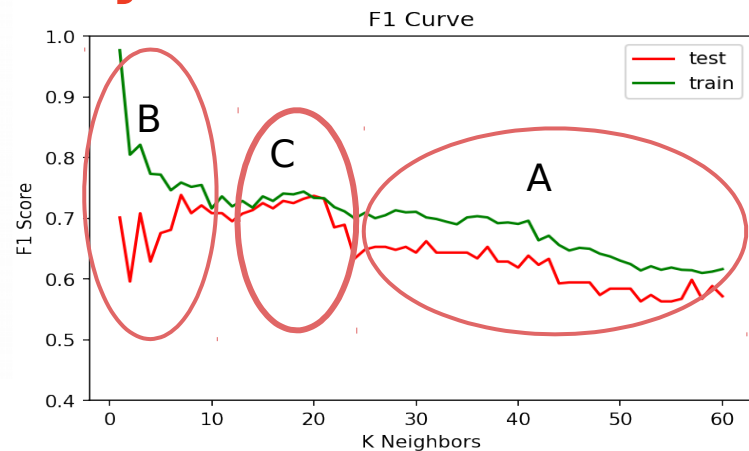


C

Conceptual Learning

Class test ~92%
Test ~89%

Best-fit



Bias and Variance Summary:



A

Not interested in learning

Class test ~50%
Test ~47%

Under-fit/ less learning



B

Memorizing the lessons

Class test ~98%
Test ~69%

Over-fit/ Memorizing

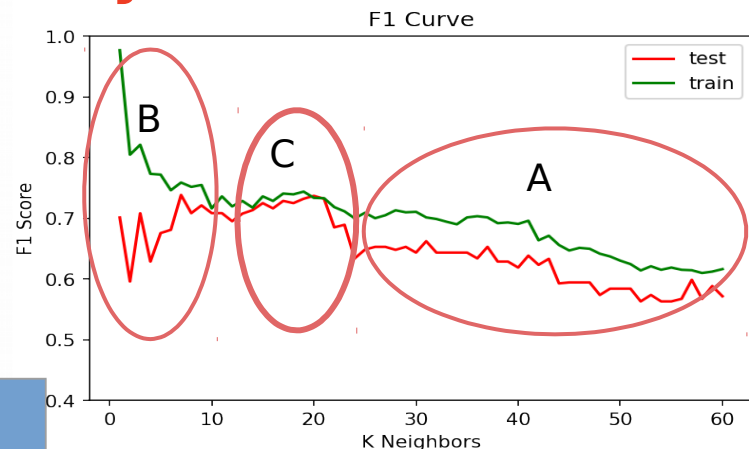


C

Conceptual Learning

Class test ~92%
Test ~89%

Best-fit



Error	High Bias	High Variance	Optimally in between
Fit	Underfit	Overfit	Bestfit
k-range	$21 < k$	$K < 11$	$12 < k < 21$
Complexity	Low Complexity	High Complexity	Optimum Complexity

Bias and Variance Summary:



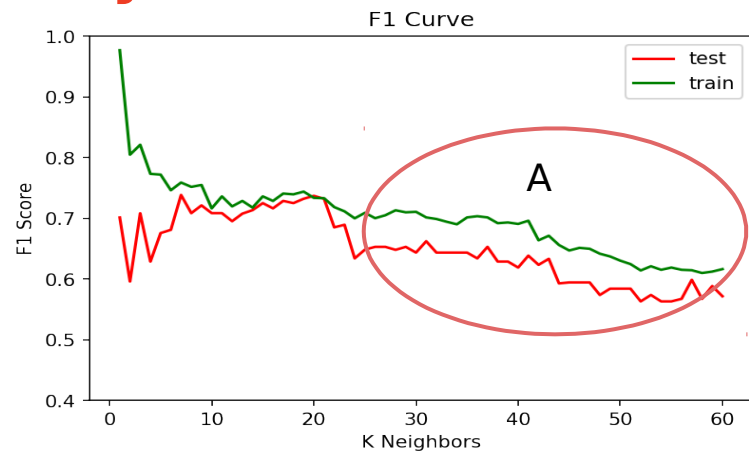
A

Not interested in learning

Class test ~50%
Test ~47%

Under-fit/ less learning

Error	High Bias
Fit	Underfit
k-range	$21 < k$
Complexity	Low Complexity



Bias and Variance Summary:



A

Not interested in learning

Class test ~50%
Test ~47%

Under-fit/ less learning



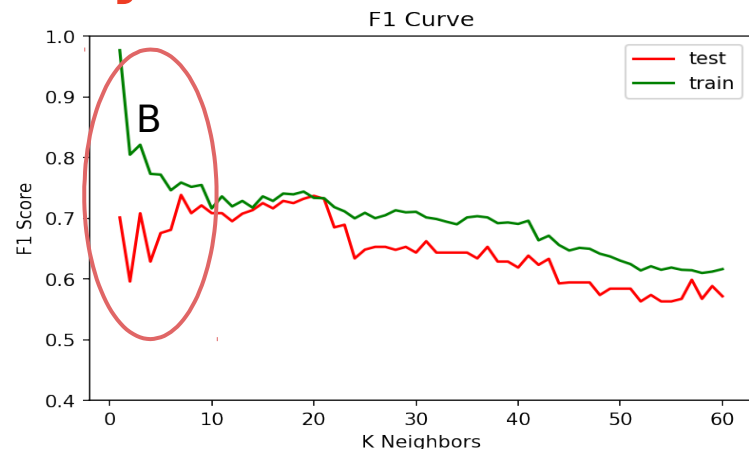
B

Memorizing the lessons

Class test ~98%
Test ~69%

Over-fit/ Memorizing

Error	High Bias	High Variance
Fit	Underfit	Overfit
k-range	$21 < k$	$K < 10$
Complexity	Low Complexity	High Complexity



Bias and Variance Summary:



A

Not interested in learning

Class test ~50%
Test ~47%

Under-fit/ less learning



B

Memorizing the lessons

Class test ~98%
Test ~69%

Over-fit/ Memorizing

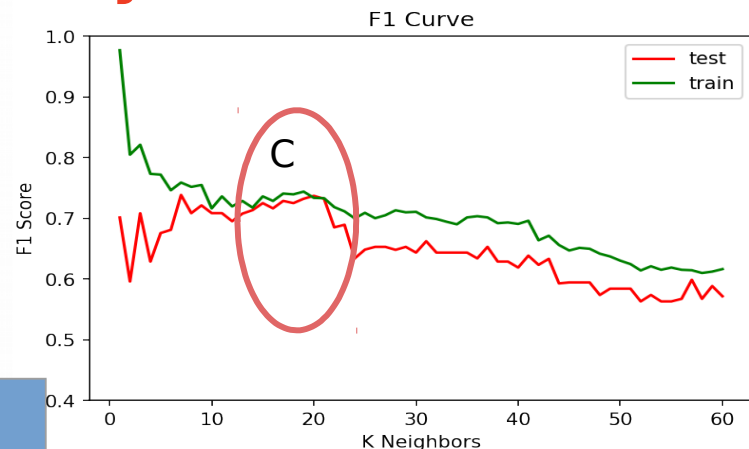


C

Conceptual Learning

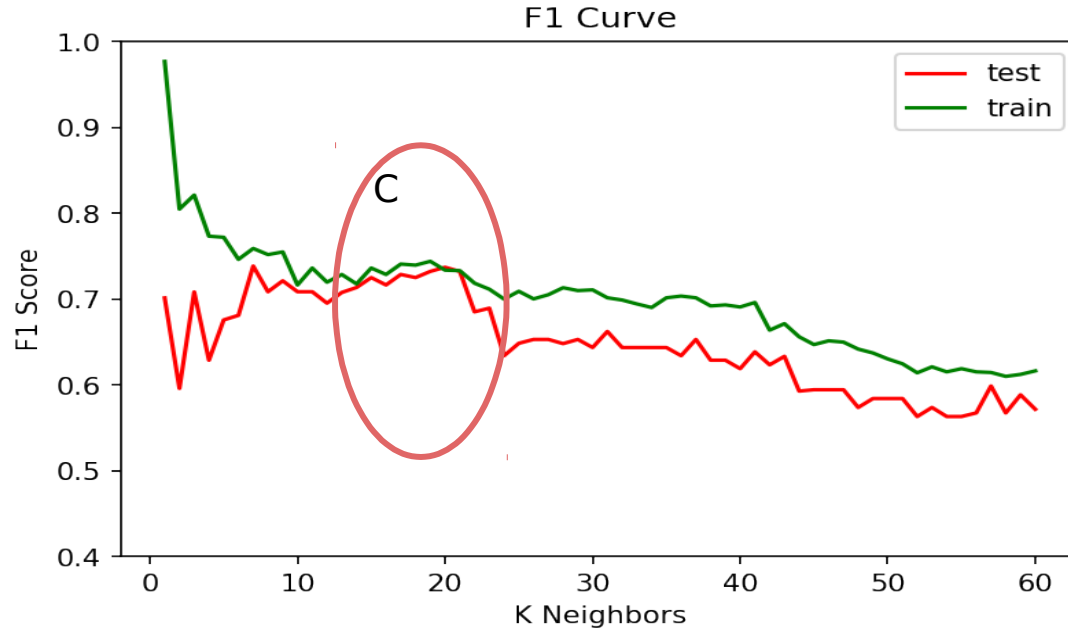
Class test ~92%
Test ~89%

Best-fit



Error	High Bias	High Variance	Optimally in between
Fit	Underfit	Overfit	Bestfit
k-range	$23 < k$	$K < 10$	$12 < k < 21$
Complexity	Low Complexity	High Complexity	Optimum Complexity

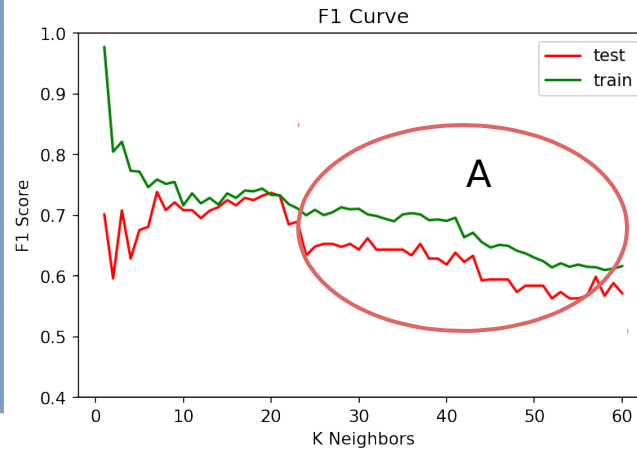
Bias and Variance Summary:



Bias

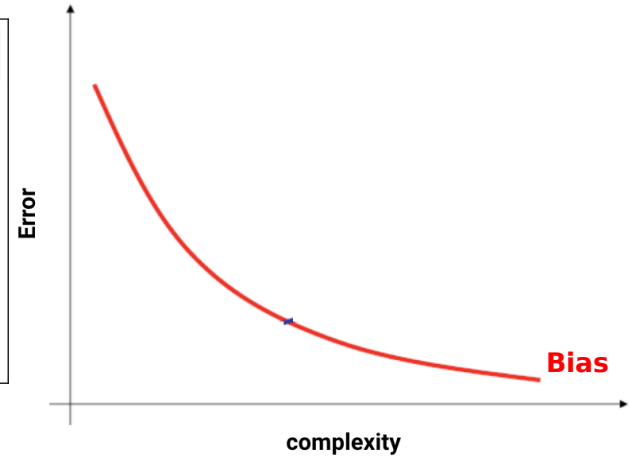
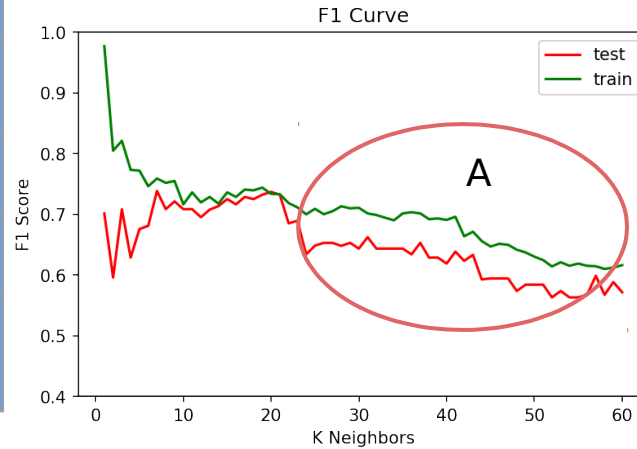
Bias

Student =>	A
Error	High Bias
Fit	Underfit
k-range	$21 < k$
Complexity	Low Complexity



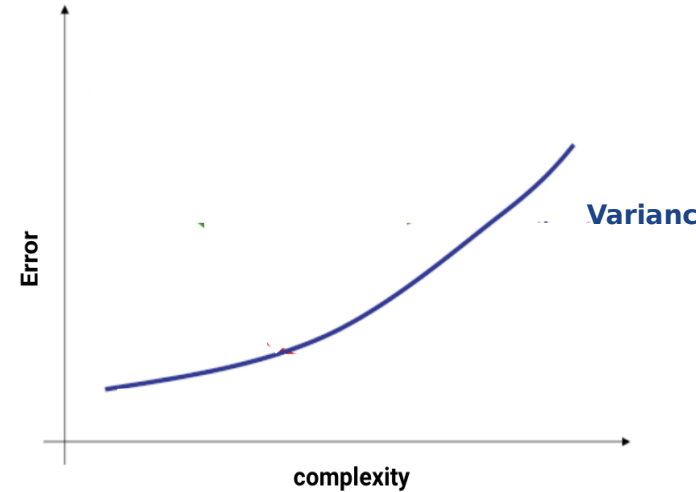
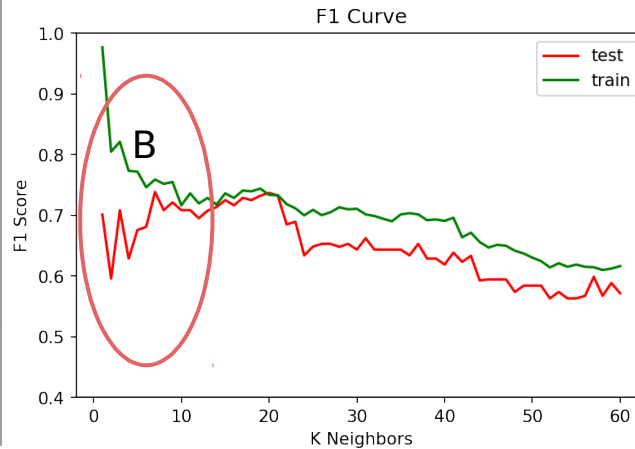
Bias

Student =>	A
Error	High Bias
Fit	Underfit
k-range	$21 < k$
Complexity	Low Complexity



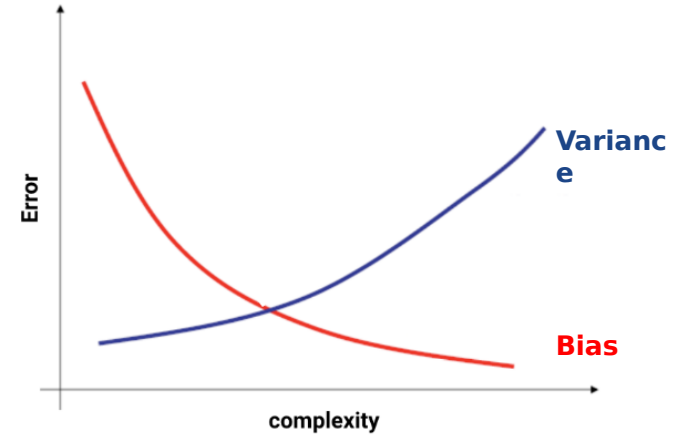
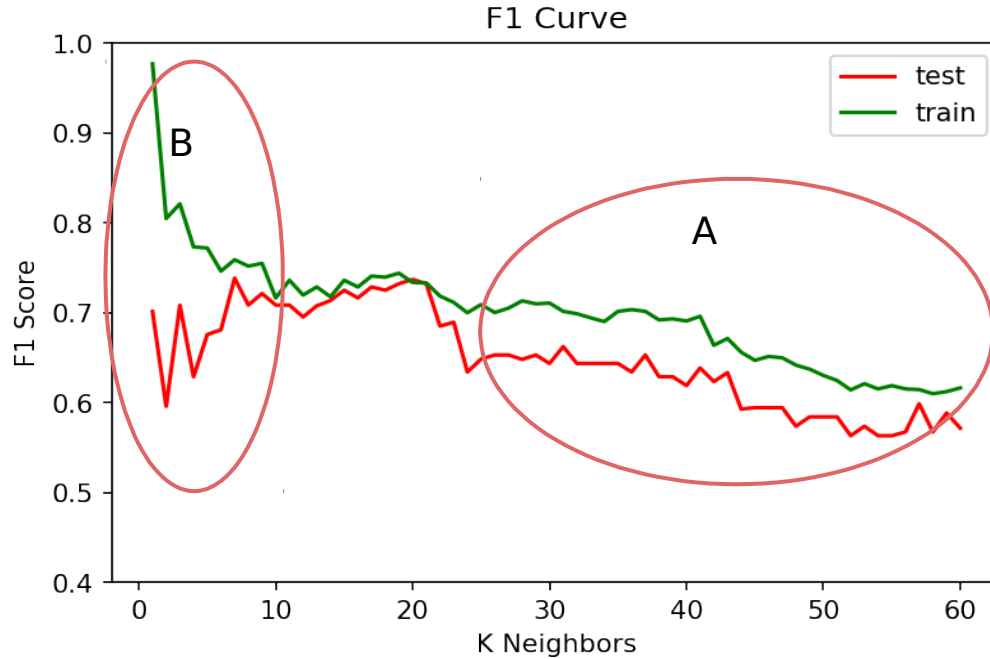
Variance

Student =>	B
Error	High Variance
Fit	Overfit
k-range	$K < 12$
Complexity	High Complexity



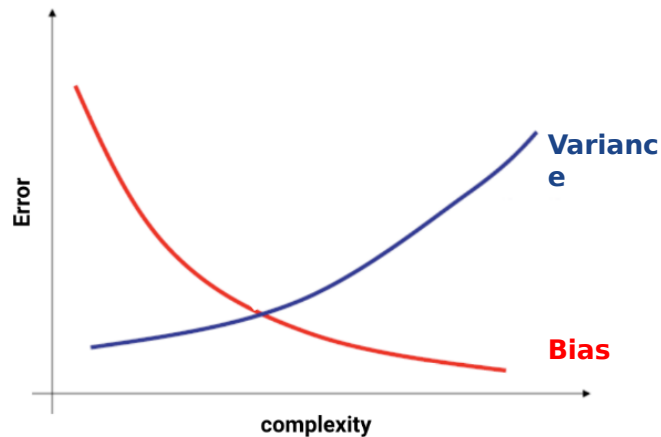
Bias and Variance Relation

Bias and Variance Relation



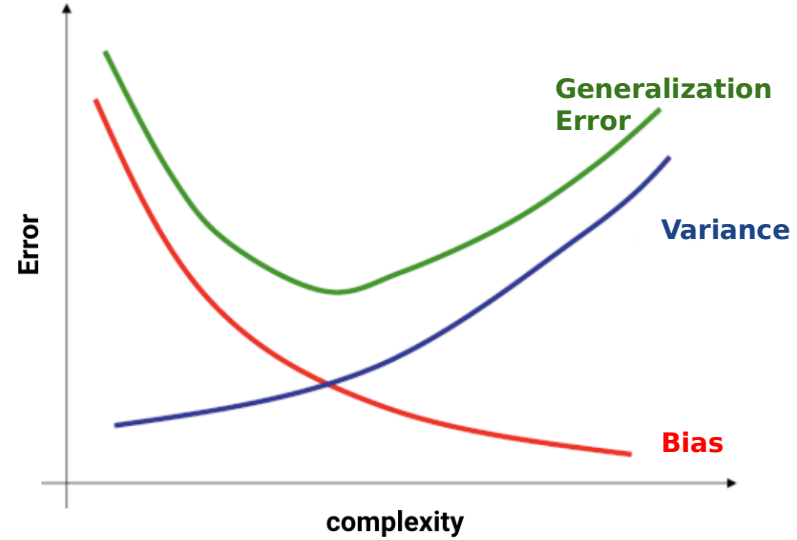
Bias and Variance Tradeoff

- Strong tradeoff
- High Variance \Rightarrow Low Bias
High Bias \Rightarrow Low Variance



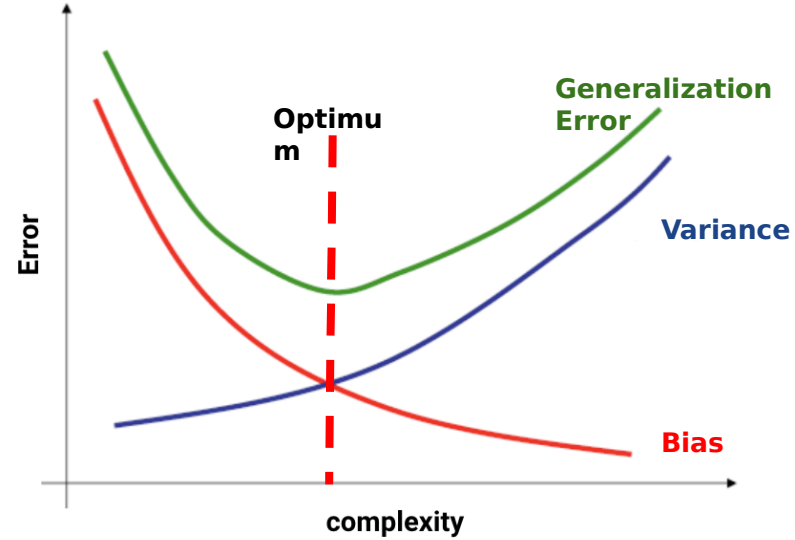
Bias and Variance Tradeoff

- Strong tradeoff
- High Variance \Rightarrow Low Bias
High Bias \Rightarrow Low Variance



Bias and Variance Tradeoff

- Strong tradeoff
- High Variance \Rightarrow Low Bias
High Bias \Rightarrow Low Variance
- Optimum in between



Bias and Variance Tradeoff

- Strong tradeoff
- High Variance \Rightarrow Low Bias
High Bias \Rightarrow Low Variance
- Optimum in between
- Complexity where Bias and Variance Error are low together

