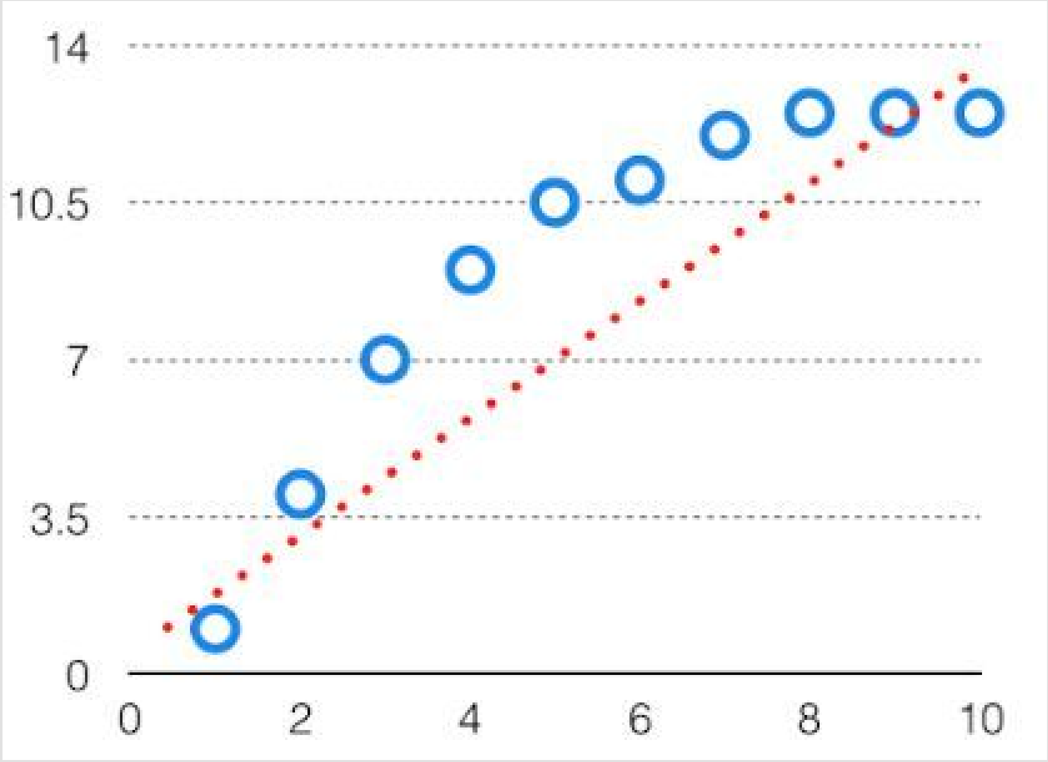
# THE BIAS VARIANCE TRADE-OFF

AVILA CHAN GABRIELA
BAAS CABAÑAS RAY
CACH ROSAS JHAIR
CASTILLO FERNANDEZ BRENDA
CHI CENTENO MARIANA
CHI GONGORA ENRIQUE

# WHAT IS THE BIAS-VARIANCE TRADE-OFF?

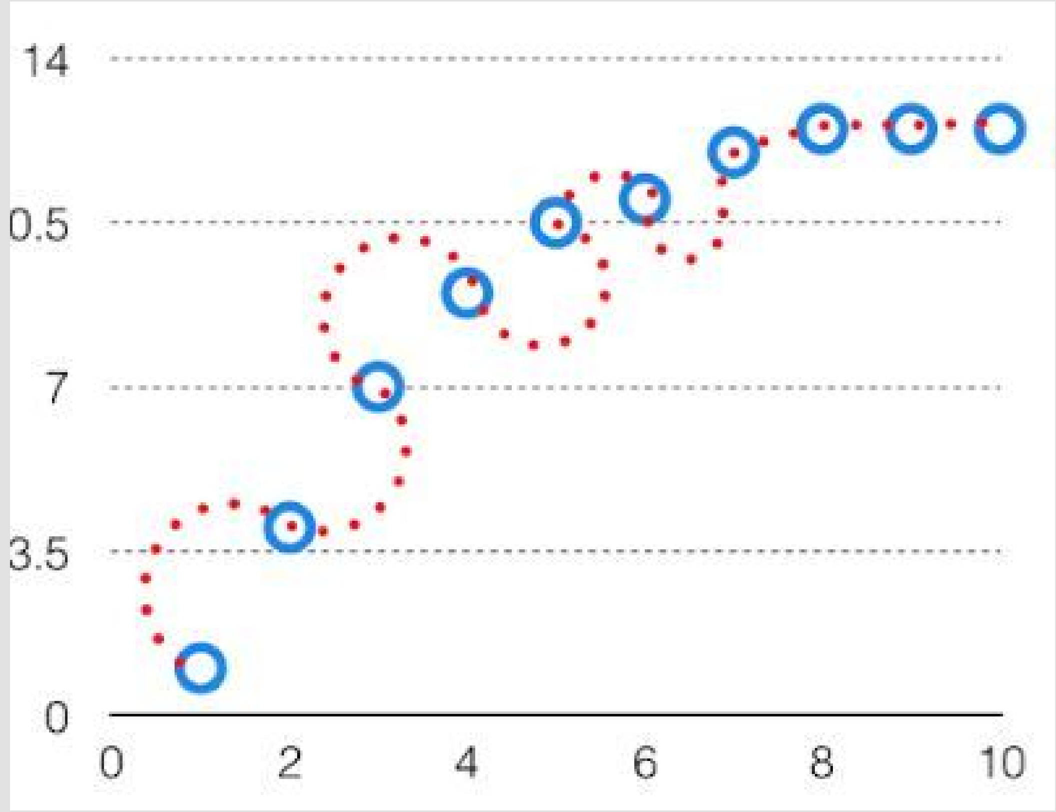
## BIAS

The bias of our model has to do with the assumptions that it makes about the data, and how well it fits to it when it is trained.




## VARIANCE

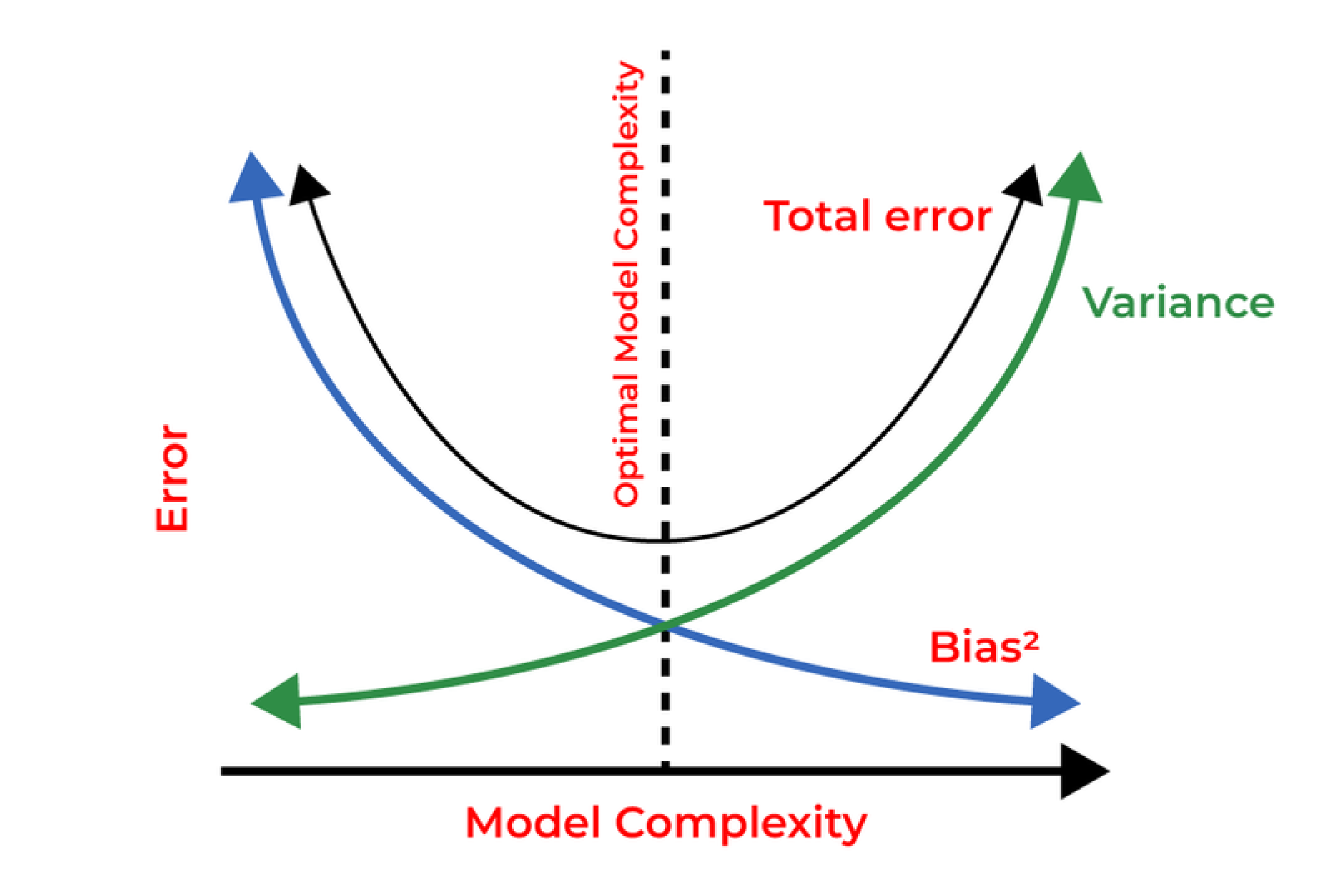
The variance of our model has to do with how it varies its results depending on the sample of data that it uses for its training.



# WHAT IS THE BIAS-VARIANCE TRADE-OFF?

The bias-variance tradeoff refers to the balance that is needed between bias and variance to build a model that can generalize well to new data.

The bias-variance trade-off in machine learning (ML) is a foundational concept that affects a supervised model's predictive performance and accuracy.



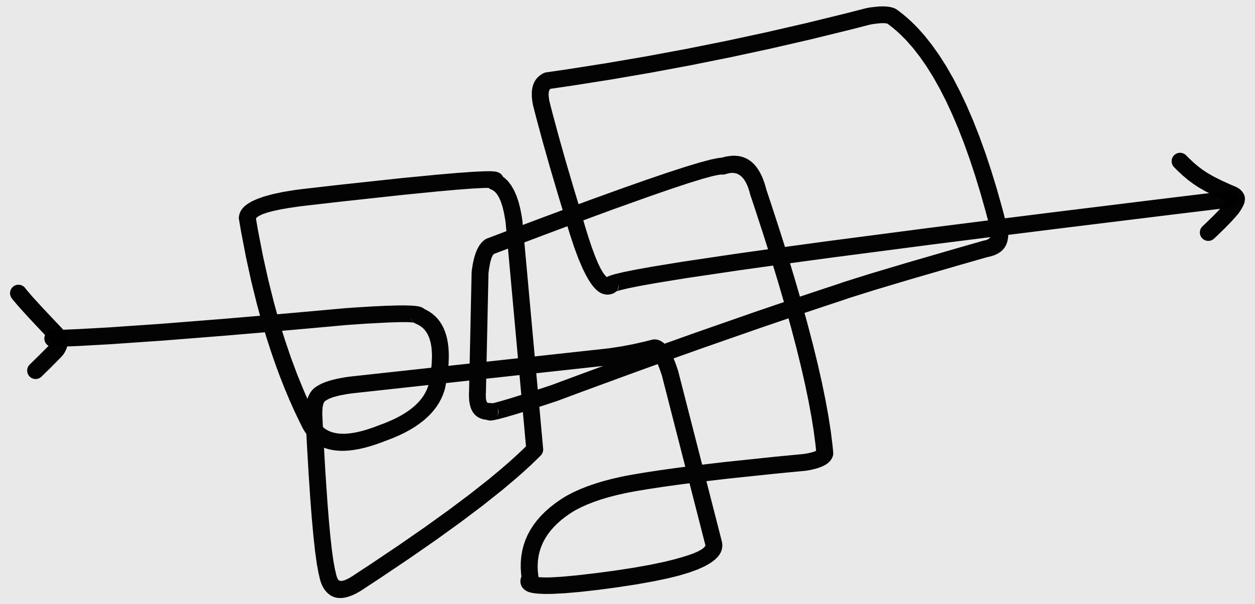
# THE EFFECTS OF HIGH BIAS

## Underfitting

A data model is unable to capture the relationship between the input and output variables accurately. It occurs when a model is too simple, which can be a result of a model needing more training time, more input features, or less regularization.

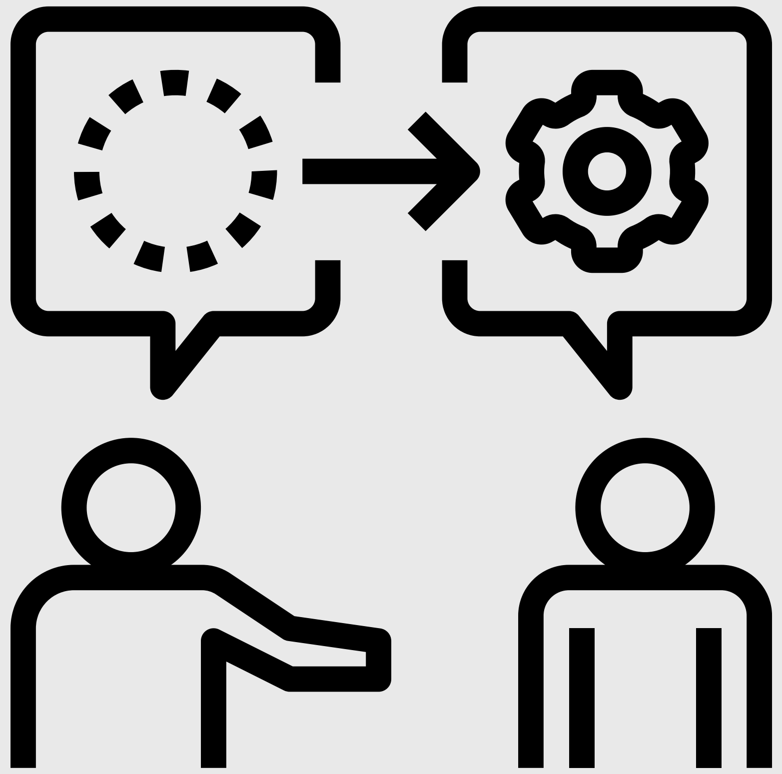
## Low Complexity

Due to the model being too simple, the biased model is unable to learn complex features of a training data, thus, making it inefficient when solving complex problems.



## Low Training Accuracy

Due to the inability to correctly process training data, the biased model shows high-training loss resulting in low-training accuracy.



# TECHNIQUES TO REDUCE BIAS IN MODELS

- BY ADDING MORE FEATURES
- BY INCREASING TRAINING ITERATIONS
- REPLACING CURRENT MODEL
- USING NON-LINEAR ALGORITHMS
- USING NON-PARAMETERIZED ALGORITHMS
- BY DECREASING REGULARIZATION
- BY USING A NEW MODEL ARCHITECTURE

# The Effects of High Variance

A HIGH VARIANCE MODEL LEADS TO OVERFITTING

INCREASE MODEL COMPLEXITIES

SENSITIVITY TO TRAINING DATA

# TECHNIQUES TO REDUCE VARIANCE IN MODELS

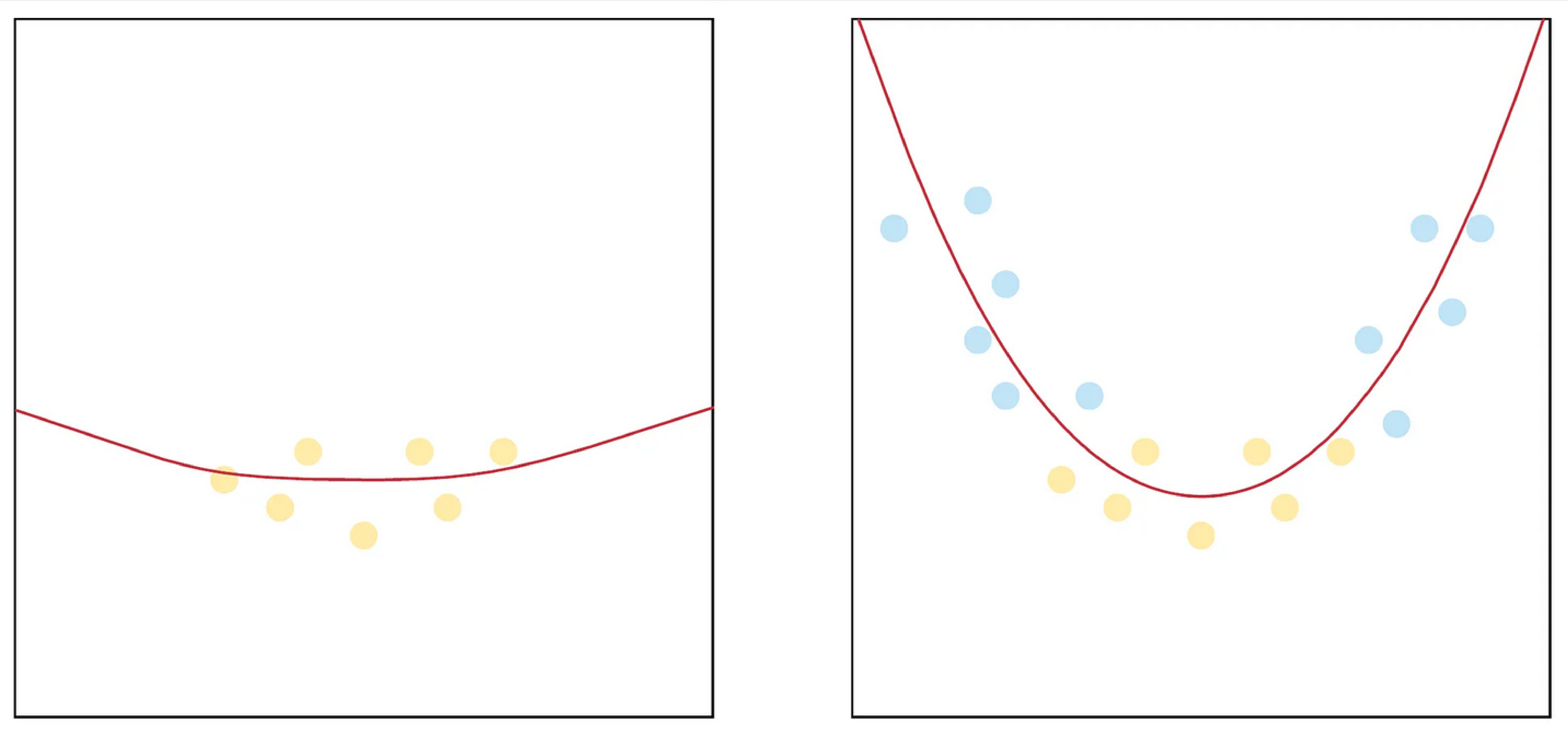

**ADD MORE DATA**

**FEATURE SELECTION**

**REGULARIZATION**

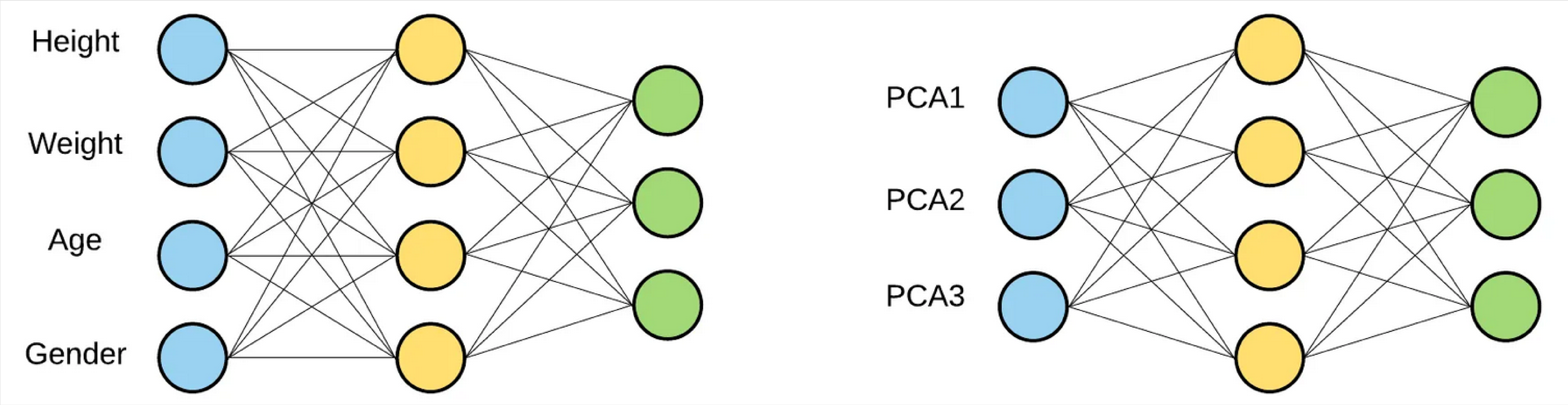
# ADD MORE DATA

By increasing the size of the dataset, the model has more examples to learn from, reducing the risk of overfitting.



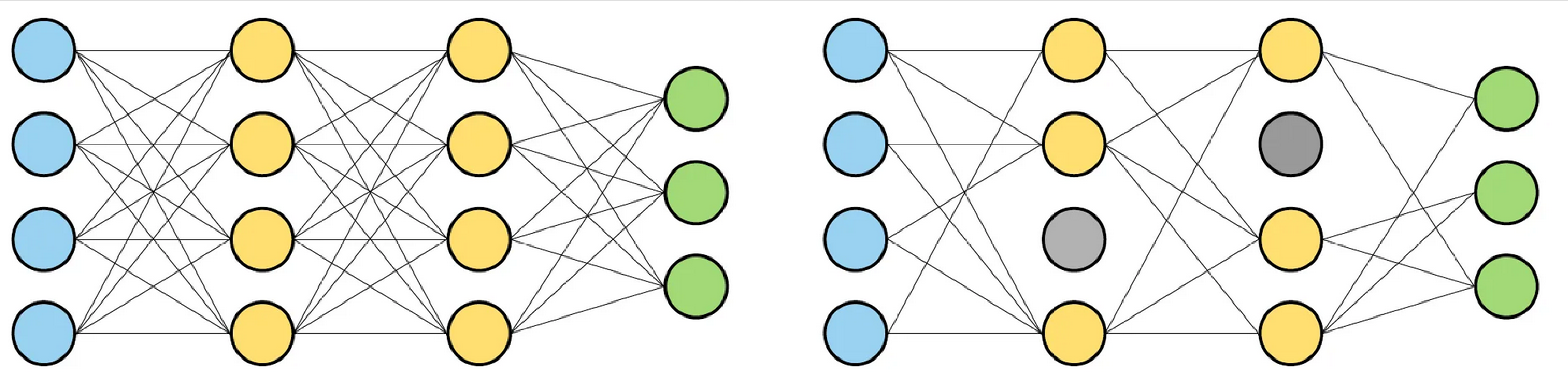
# FEATURE SELECTION

By selecting the most important features, the model has fewer opportunities to overfit the data by capturing less relevant information.



# REGULARIZATION

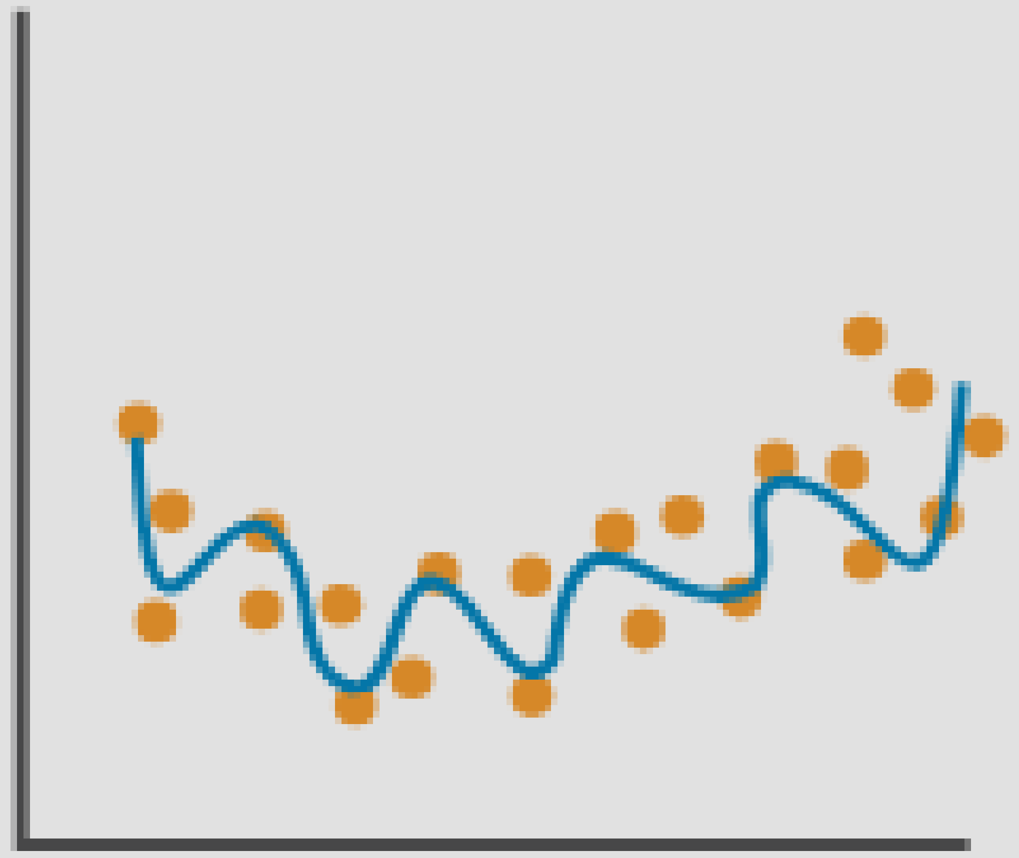
Adding penalties to models, like L1 or L2 regularization, can help avoid overfitting by constraining the model's parameter values.



# THE BIAS-VARIANCE TRADE-OFF IN ACTION

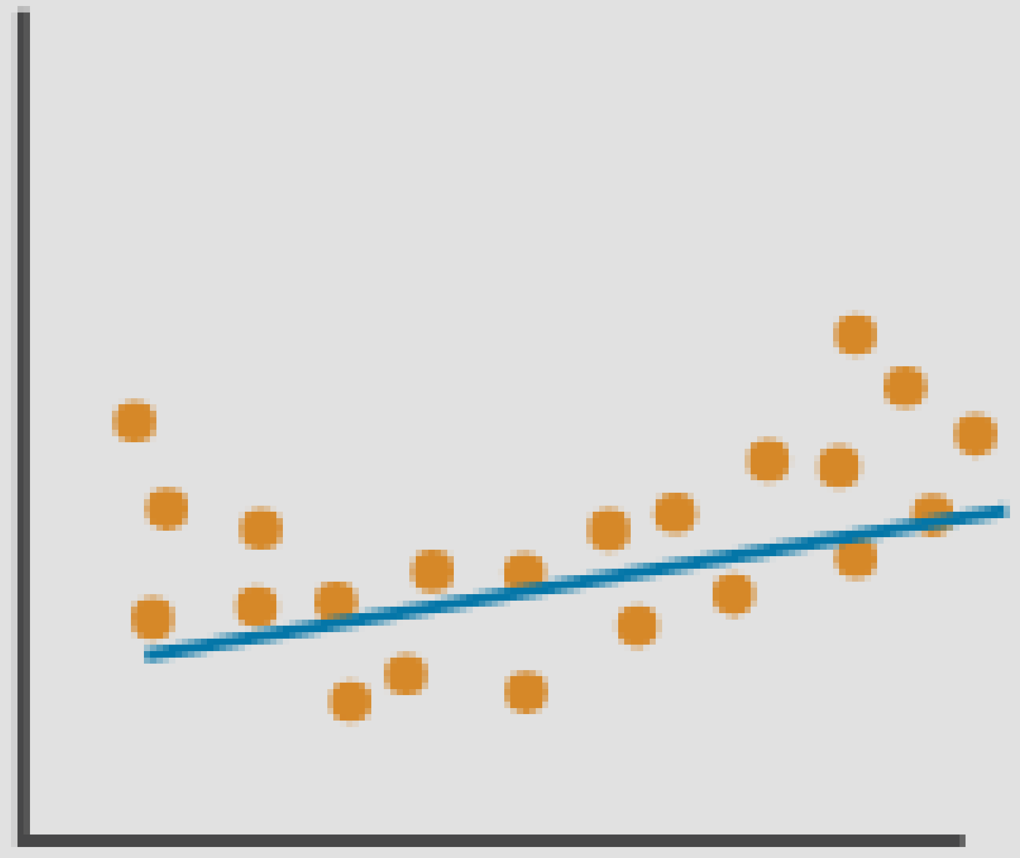
## LOW-BIAS, HIGH-VARIANCE:

Inconsistent predictions and, on average, accurate. This occurs when the model learns with a large number of parameters, leading to overfitting.



## HIGH-BIAS, LOW-VARIANCE:

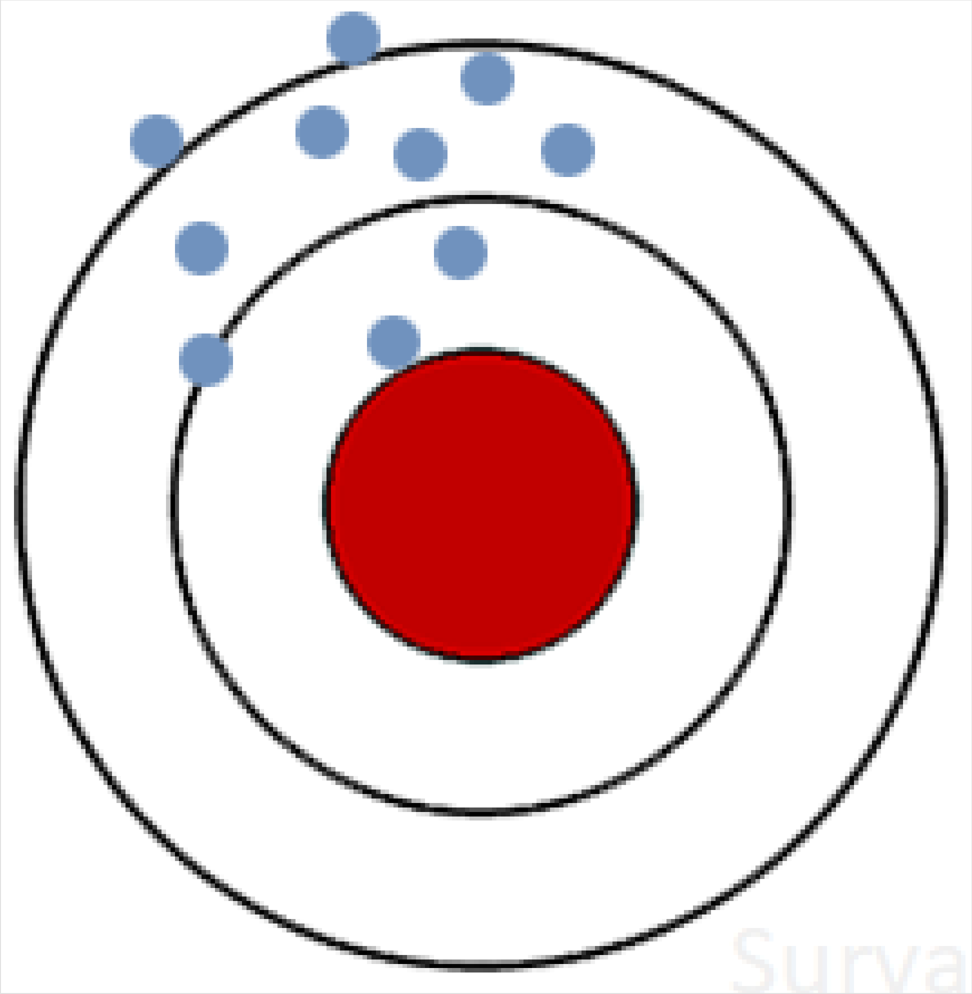
Consistent predictions but, on average, inaccurate. Occurs when the model doesn't learn well with the training dataset or uses few parameters, leading to underfitting.



# THE BIAS-VARIANCE TRADE-OFF IN ACTION

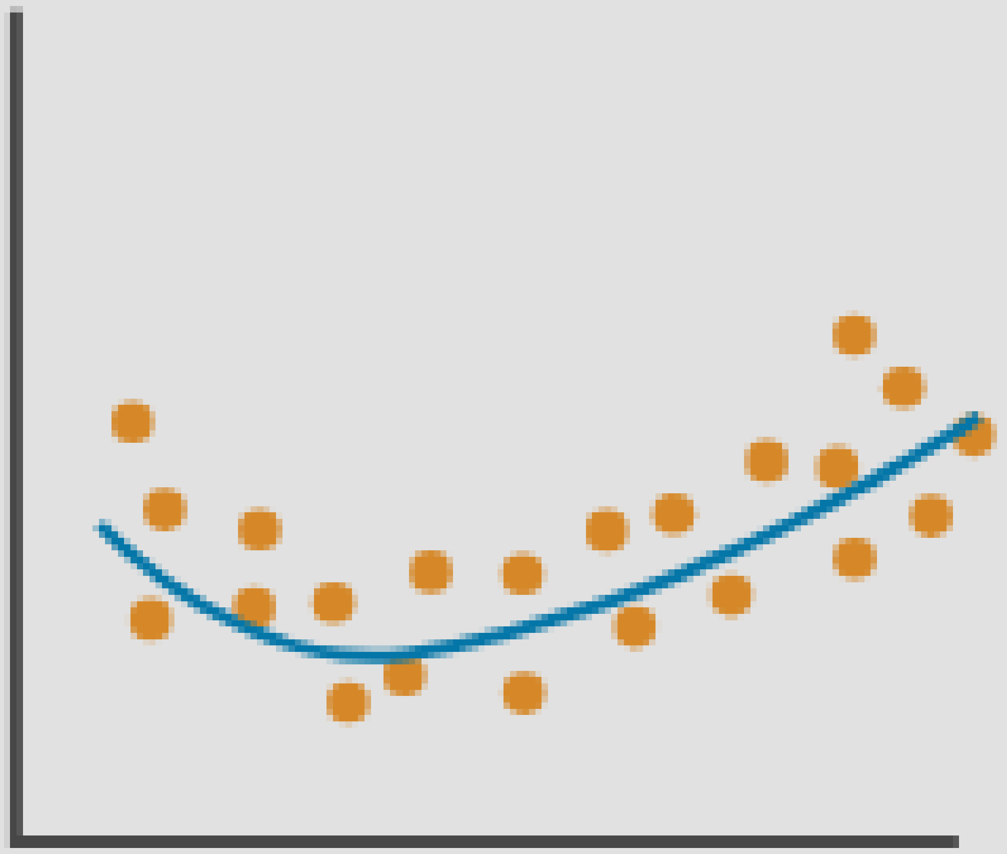
## HIGH-BIAS, HIGH-VARIANCE:

Inconsistent and, on average, inaccurate predictions. The model cannot capture underlying patterns in the data and is sensitive to changes in the training data.



## OPTIMAL TRADE-OFF:

The combination of low bias and low variance represents an ideal machine learning model that can effectively capture underlying patterns in the data and generalize to new data. However, in practice, achieving this perfect balance is extremely challenging.



# ENSEMBLE METHODS, BIAS, AND VARIANCE

## Bagging and variance

Is meant to reduce the variance without increasing the bias. This technique is especially effective where minute changes in a learner's training set lead to huge changes in the predicted output.

## Boosting and bias

Boosting is especially useful in models that exhibit underfitting. It shows how poorly a function fits the given data points. To deal with this error, we train a learner and identify where it exhibits bias errors. The observations that are wrongly classified are assigned higher weights.

# Implications and Applications

01

## Higher Accuracy

Understanding and optimizing the bias-variance trade-off can lead to more accurate models and predictions.

02

## Interpretability

By creating simpler models that still perform well, we can better understand the underlying relationships within the data.

03

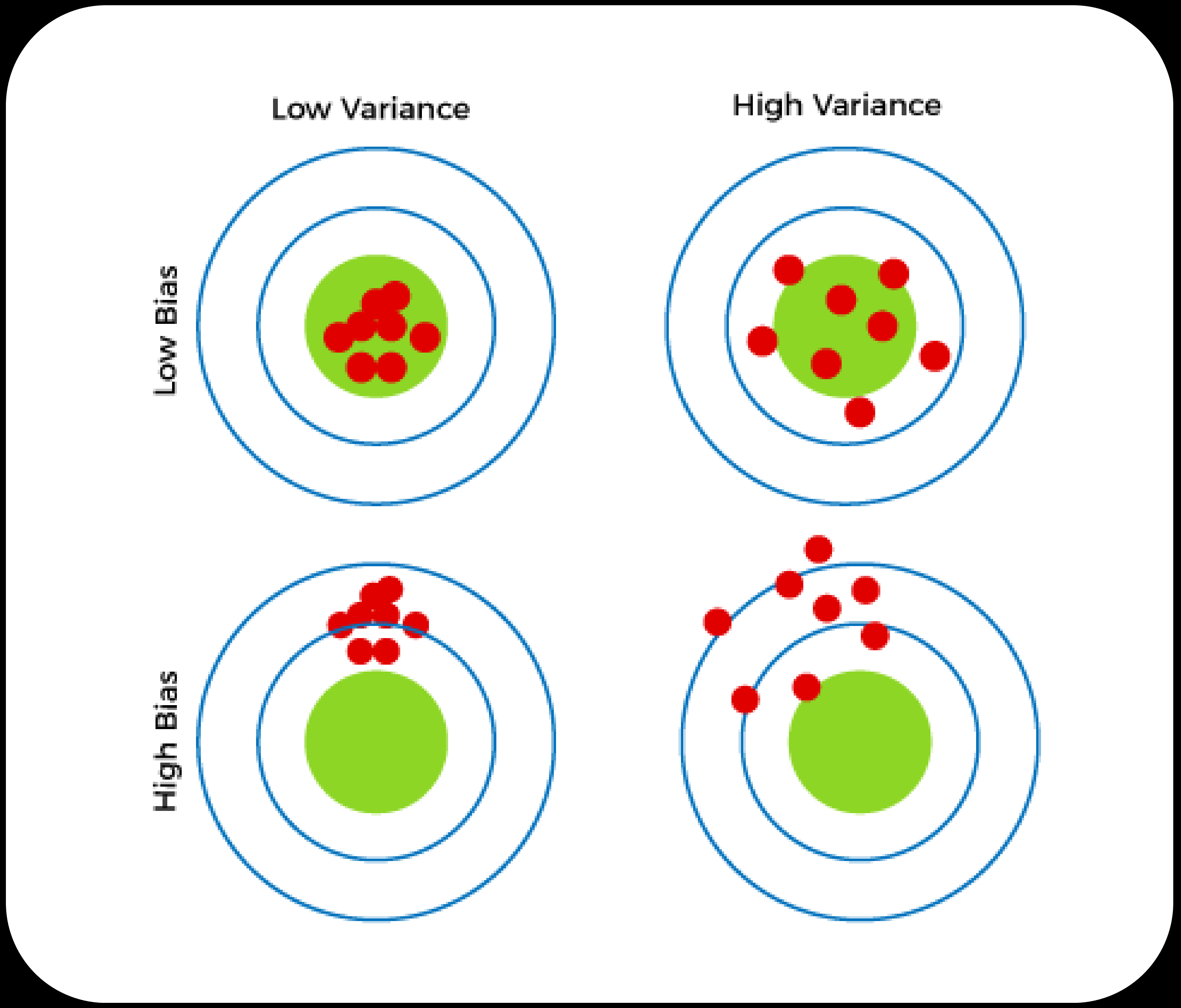
## Cost Savings

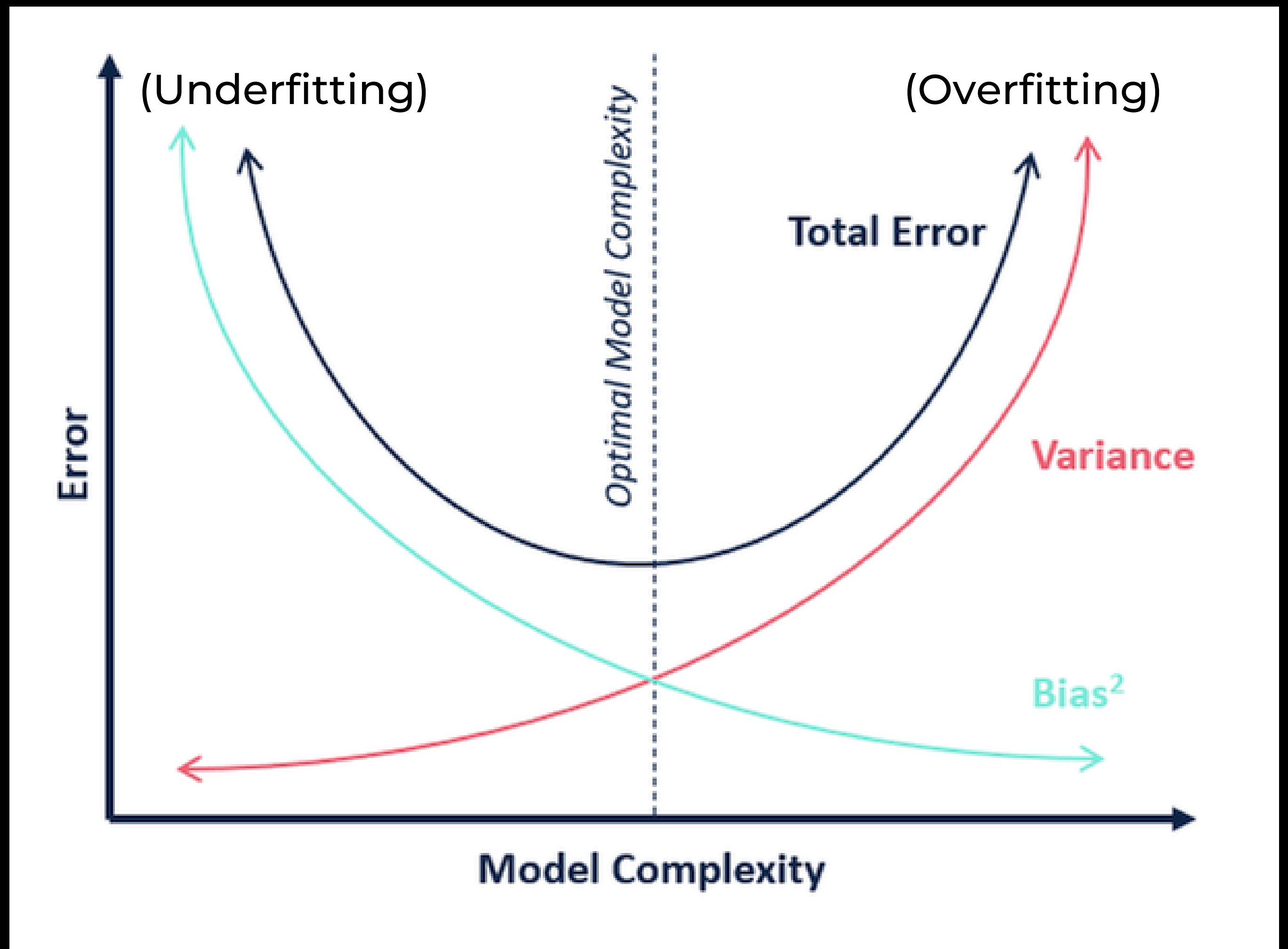
Saving time and money on unnecessary models that could cause overfitting and end up being too complex.

04

## Applications

The bias-variance trade-off can be applied to many machine learning related tasks, including natural language processing, image recognition, and fraud detection.




# EXAMPLE

## Predicting Exam Scores

1

### HIGH BIAS, LOW VARIANCE (UNDERFITTING):

Using a simple linear regression model to predict exam scores based on study hours, assuming a linear relationship. Results in systematic errors, underestimating or overestimating scores for all students, and oversimplification of student performance.

2

### BALANCED BIAS AND VARIANCE (GOOD FIT):

Choosing a more flexible model like polynomial regression, carefully tuning complexity, and regularization to balance bias and variance. Accurate predictions based on study hours, generalizing to different student populations.

3

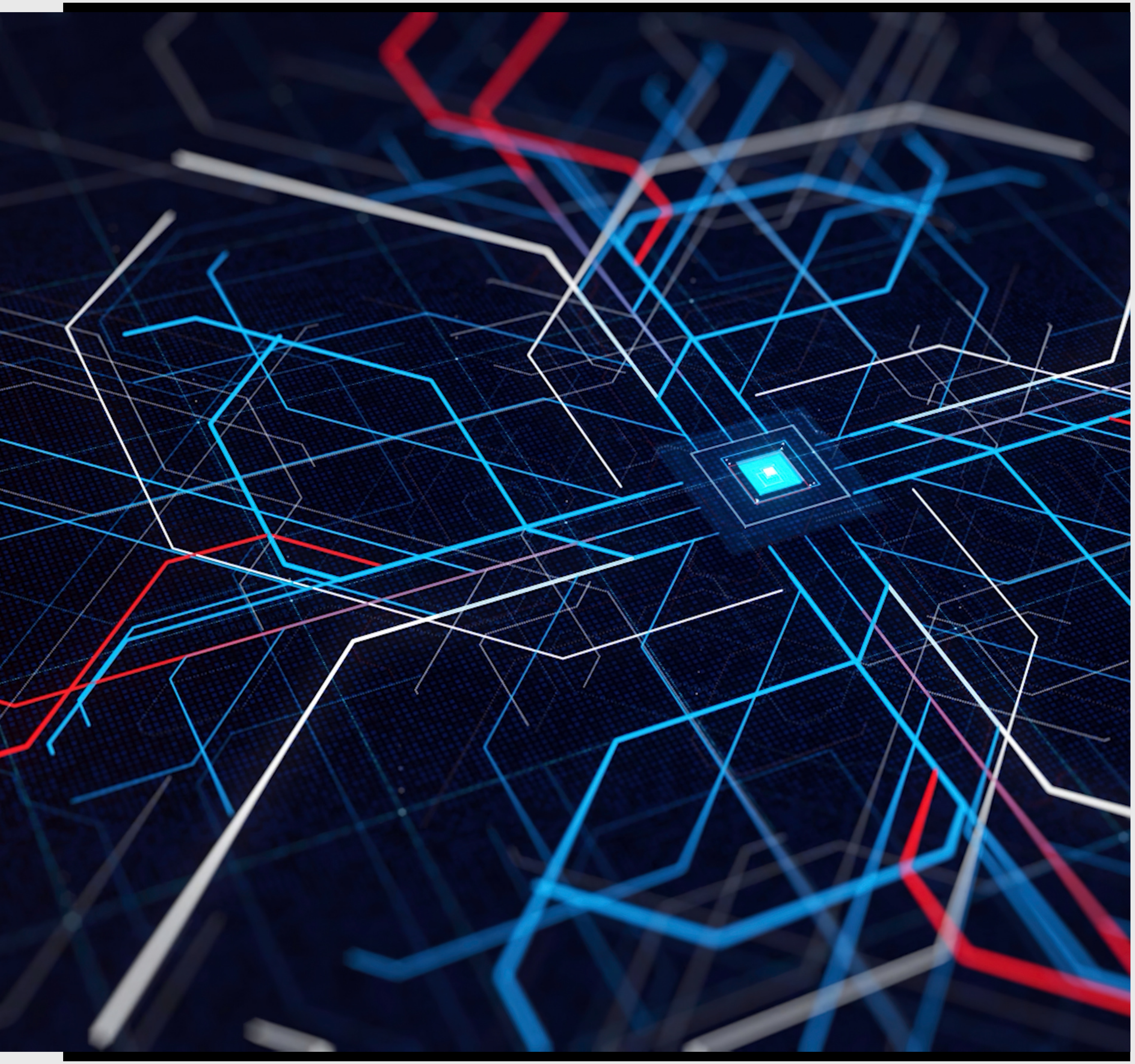
### LOW BIAS, HIGH VARIANCE (OVERFITTING):

Implementing an overly complex deep neural network, capturing noise in the training data, and leading to excellent performance on training data but poor generalization to new data due to sensitivity to fluctuations.

# EXAMPLE

## Predicting Exam Scores

THE BIAS-VARIANCE TRADEOFF TEACHES US THE IMPORTANCE OF CHOOSING THE RIGHT MODEL COMPLEXITY LEVEL. IT'S ABOUT FINDING A BALANCE: TOO SIMPLE A MODEL WON'T CAPTURE NUANCES, WHILE AN OVERLY COMPLEX ONE BECOMES UNRELIABLE. STRIKING THIS BALANCE IS CRUCIAL FOR MAKING ACCURATE PREDICTIONS, LIKE FORECASTING NEW STUDENTS' EXAM SCORES BASED ON THEIR STUDY HOURS.



# CONCLUSION

The bias-variance trade-off is a crucial concept in machine learning, and finding the optimal balance is necessary for developing accurate models. By understanding the implications and applications of the trade-off, we can create models that generalize well to new data and produce accurate predictions.

# REFERENCES


- Ayuya, C. (n.d.). Ensemble learning on bias and variance. Retrieved September 21, 2023, from Engineering Education (EngEd) Program | Section website: https://www.section.io/engineering-education/ensemble-bias-var/
- Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proceedings of the National Academy of Sciences, 116(32), 15849-15854.
- Bias and variance in machine Learning - Javatpoint. (s. f.). www.javatpoint.com. https://www.javatpoint.com/bias-and-variance-in-machine-learning
- GeeksforGeeks. (2023). Bias and variance in machine learning. GeeksforGeeks. https://www.geeksforgeeks.org/bias-vs-variance-in-machine-learning/
- Guan, X., & Burton, H. (2022). Bias-variance tradeoff in machine learning: Theoretical formulation and implications to structural engineering applications. In Structures (Vol. 46, pp. 17–30). Elsevier BV. https://doi.org/10.1016/j.istruc.2022.10.004
- Dar, Y., Muthukumar, V., & Baraniuk, R. G. (2021). A farewell to the bias-variance tradeoff? an overview of the theory of overparameterized machine learning. arXiv preprint arXiv:2109.02355.
- Doroudi, S. (2020). The bias-variance tradeoff: How data science can inform educational debates. AERA open, 6(4), 2332858420977208.
- Just Understanding Data. (2023, 19 abril). Introduction to the Bias-Variance Trade-Off in Machine Learning - Just Understanding Data. Understanding Data. https://understandingdata.com/posts/introduction-to-the-bias-variance-trade-off-in-machine-learning/
- Pham, K. (2023, 12 mayo). Overfitting, generalization, & the Bias-Variance tradeoff. Medium. https://medium.com/@khang.pham.exxact/overfitting-generalization-the-bias-variance-tradeoff-5800f8c2200
- Sanz, F. (2021). Entiende de una vez qué es el tradeoff Bias-Variance. The Machine Learners. https://www.themachinelearners.com/tradeoff-bias-variance/

# REFERENCES


- Thorn, J. (2021, 13 diciembre). Quick Bias/Variance Trade-Off - towards data science. Medium. https://towardsdatascience.com/quick-bias-variance-trade-off-d4895b126b08
- What is the tradeoff between Bias and Variance? (n.d.). Retrieved September 21, 2023, from Educative website: https://www.educative.io/answers/what-is-the-tradeoff-between-bias-and-variance
- Elliott, S. (2022, February 18). Overcoming bias & variance in machine learning. Retrieved September 21, 2023, from Towards Data Science website: https://towardsdatascience.com/overcoming-bias-variance-in-machine-learning-31169dc649ed
- Bias and Variance in Machine Learning. (n.d.). Retrieved September 21, 2023, from www.javatpoint.com website: https://www.javatpoint.com/bias-and-variance-in-machine-learning
- (N.d.). Retrieved September 21, 2023, from Researchgate.net website: https://www.researchgate.net/figure/As-model-complexity-increases-performance-on-the-data-used-to-build-the-model-training_fig1_220741326#:~:text=by%20Jonathan%20Shapiro-,As%20model%20complexity%20increases%2C%20performance%20on%20the%20data%20used%20to,This%20is%20called%20overfitting.
- Twin, A. (2010, August 14). Understanding overfitting and how to prevent it. Retrieved September 21, 2023, from Investopedia website: https://www.investopedia.com/terms/o/overfitting.asp