UNIVERSIDAD POLITECNICA DE YUCATÁN

Machine Learning

TEACHER: Victor Ortiz

STUDENT:

AXEL ESDRAS DE FLORES ISLAS

Robotics GROUP 9A

UNIVERSIDAD POLITÉCNICA DE YUCATÁN BIS Universities

Of course! Here's the simplified explanation in English:

1. Overfitting and Underfitting:

Overfitting:

When a learning model focuses too much on training data and learns weird things that aren't helpful. It works great with that data but fails with new data.

Imagine a basketball player who spends hours practicing three-point shots on his local court. He knows every imperfection in the floor and can make almost every shot there. But when he goes to play on a different court in a real game, his skills are not as impressive because he got too used to his local court.

Underfitting:

Happens when a model is too simple and can't handle training data properly. It performs poorly with both known and new data.

On the other hand, think of a player who hardly practices and doesn't take training seriously. When the game arrives, they can't make simple shots or dribble correctly. Here, underfitting would be like not having prepared enough.

2. Outliers:

These are unusual data points that stand out and can confuse analysis or predictions.

Let's say you're analyzing the scoring statistics of a basketball team. Most players score around 10-20 points per game, but there's a player who occasionally scores more than 50 points in a single game. That player is an outlier in your dataset because their performances are significantly different from those of their teammates.

3. Solutions for Overfitting, Underfitting, and Outliers:

Overfitting:

Simplify your model.

Set rules to prevent the model from going crazy.

Test the model in different situations.

If a basketball team is overtrained in a specific tactic, like full-court press, they may become vulnerable to teams with different strategies. To address it, the coach could vary their tactical approach based on the opponent and balance their playing style.

Underfitting:

Make the model smarter or use more information.

Get more data if needed.

Make the model more useful.

If a team doesn't practice enough and struggles in every game, they could improve their performance by increasing practice time, enhancing individual skills, and working together as a cohesive team.

Outliers:

Find and fix unusual data or get rid of it if it's problematic.
Use models that aren't too sensitive to unusual data.
If in a basketball team's points-per-game statistics, the outlier player who scores more than 50 points negatively impacts the analysis, the coach might consider analyzing the team's performance by excluding those exceptional games to get a more representative picture of the overall performance.
Modify the data to reduce the impact of unusual values.

4. Dimensionality Problem:
When you have too many features in your data, and it becomes challenging to manage.

5. Dimensionality Reduction:
Reducing the number of features in your data to simplify it and make it easier to handle.

6. Bias-Variance Trade-Off:
It's like finding a balance between being very stubborn in your beliefs (bias) and changing your mind all the time (variance). You need to find the middle ground to make accurate decisions.

https://www.nrigroupindia.com/e-book/Introduction%20to%20Machine%20Learning%20with%20Python%20(%20PDFDrive.com%20)-min.pdf