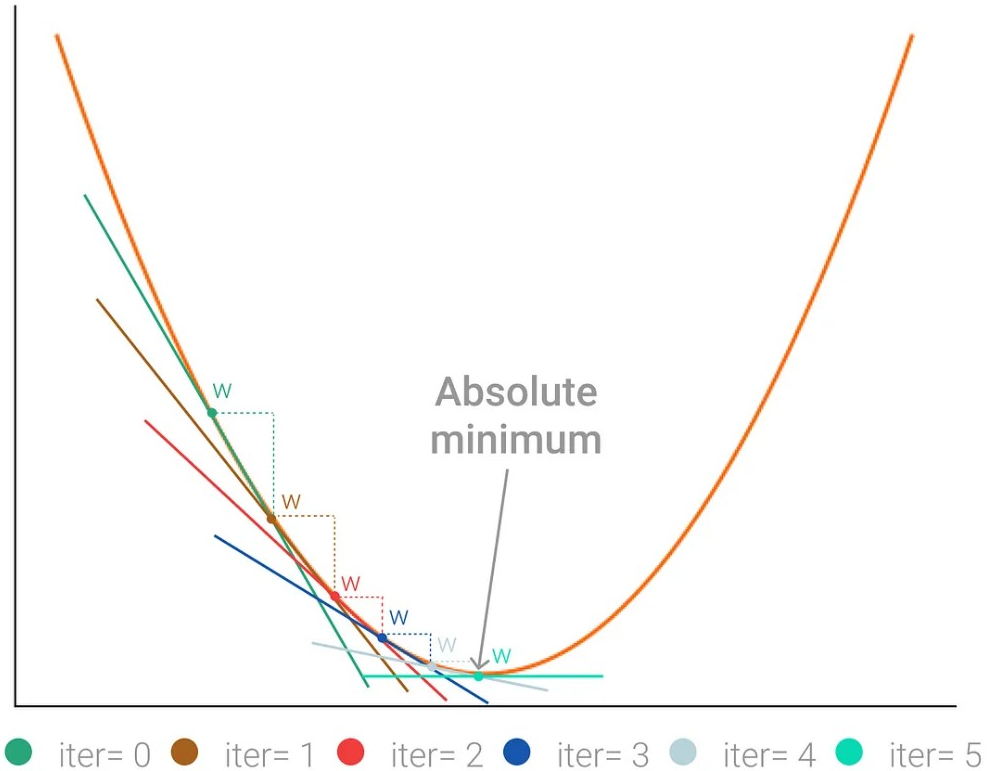
SWB-U Lagos

Additional Slides for presentation on Optimization in Machine Learning

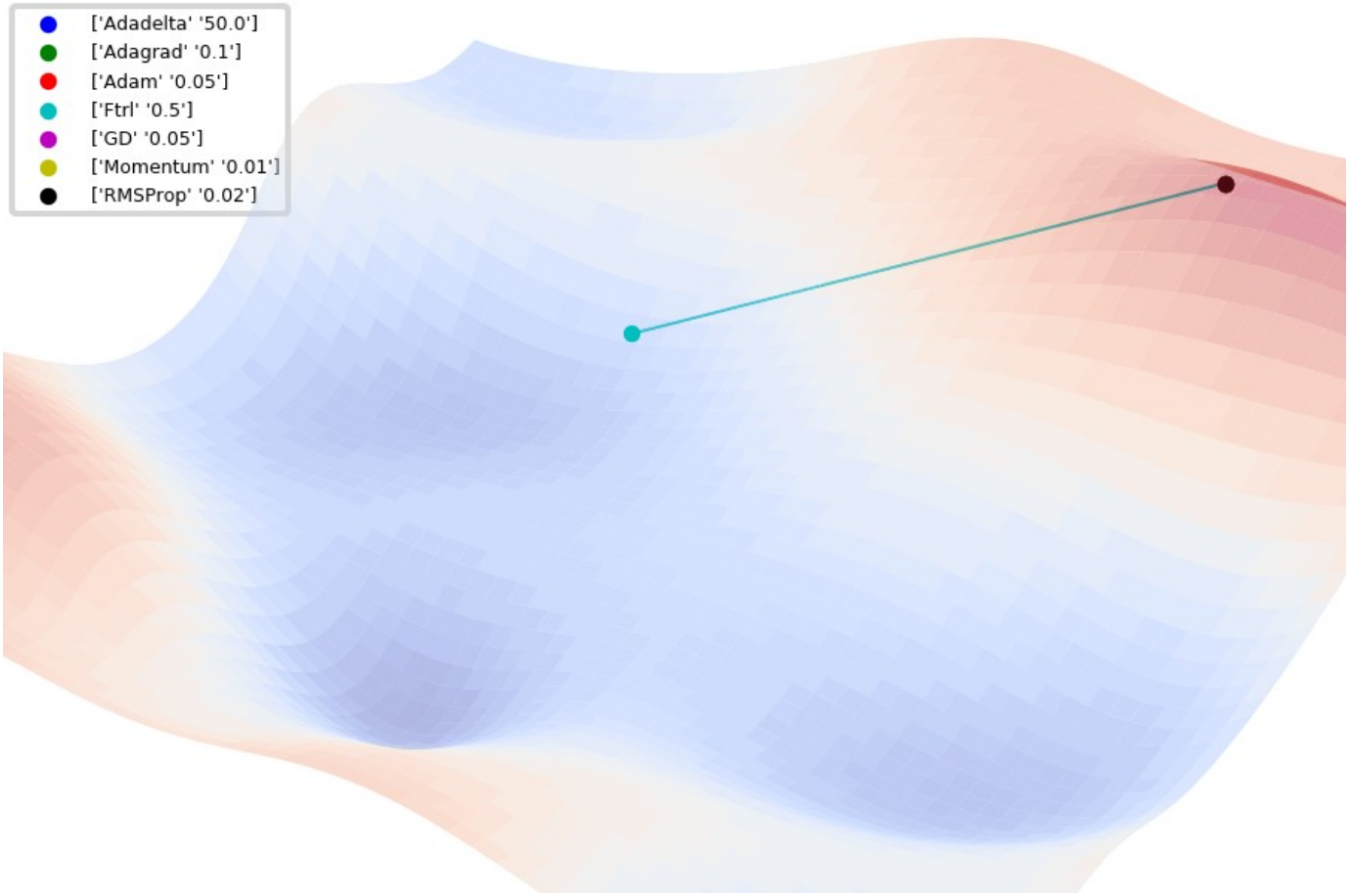
Ernesto L. Garcia C., Phd

Gradient Descent

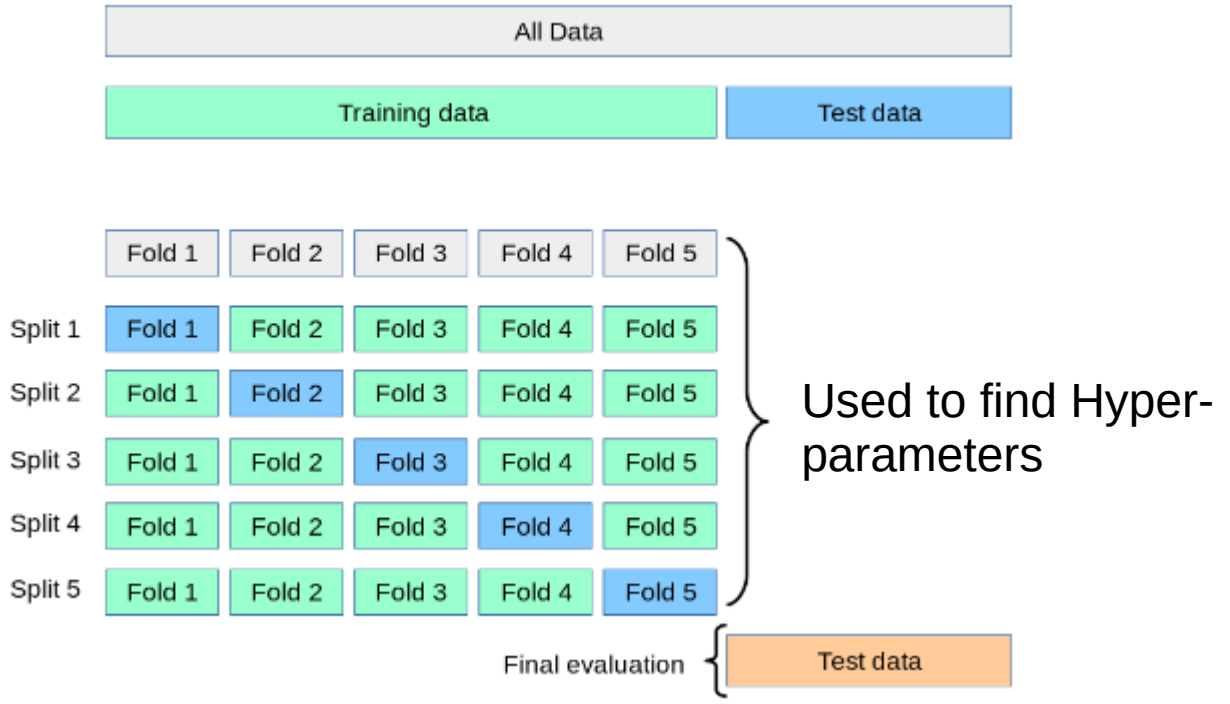


https://towardsdatascience.com/deep-learning-optimizers-436171c9e23f

SGD Visualizations

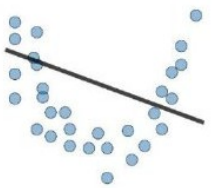
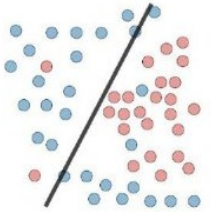
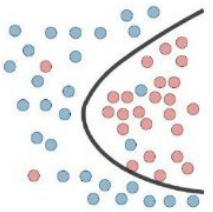
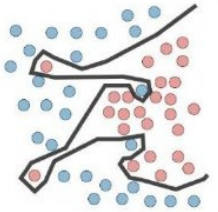
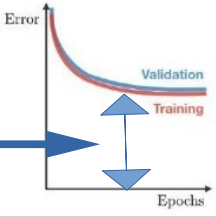
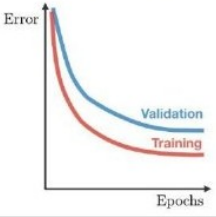
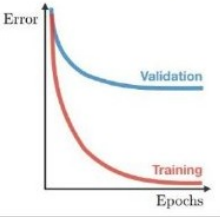


Cross Validation



Folds in blue are validation sets

Fitting types in ML

	Underfitting	Just right	Overfitting
Symptoms	- High training error - Training error close to test error - High bias	- Training error slightly lower than test error	- Low training error - Training error much lower than test error - High variance
Regression			
Classification			
Loss curves			
Remedies	- Complexify model - Add more features - Train longer		- Regularize - Get more data

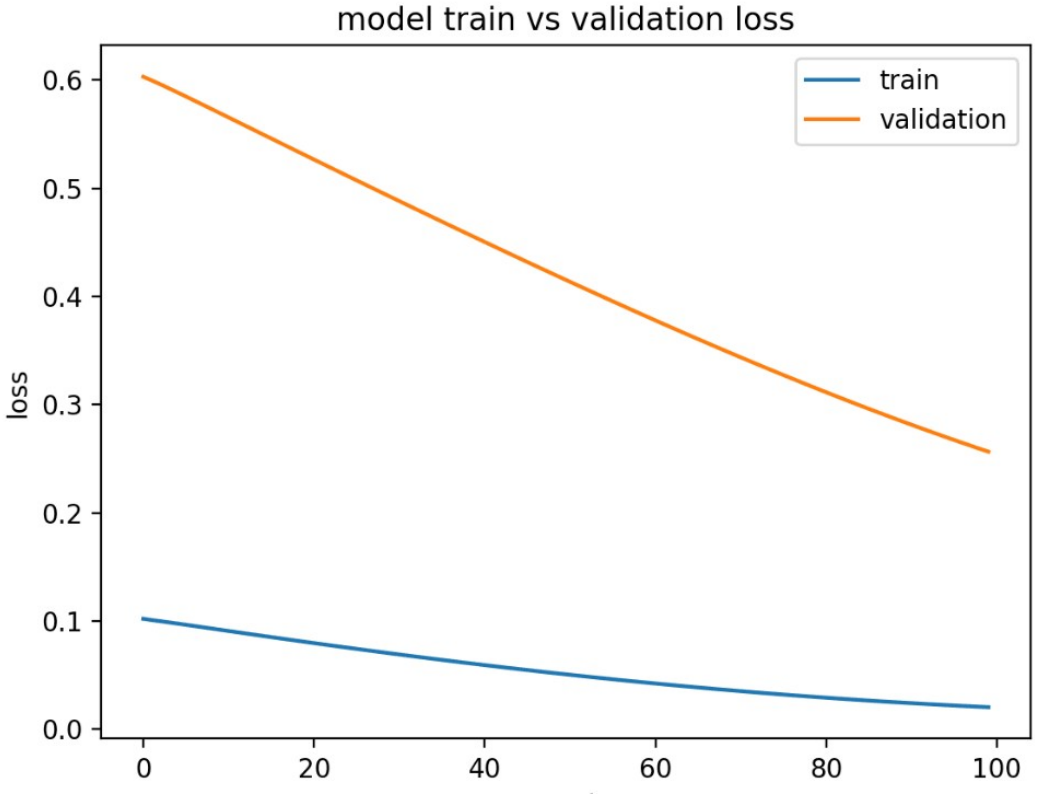
GAP

Source: https://i.pinimg.com/originals/72/e2/22/72e222c1542539754df1d914cb671bd7.png

Underfit Model Example 1

Underfit is detected in the training curve: The model is unable to learn from the data set.
Increase model complexity

Underfit Model Example 2



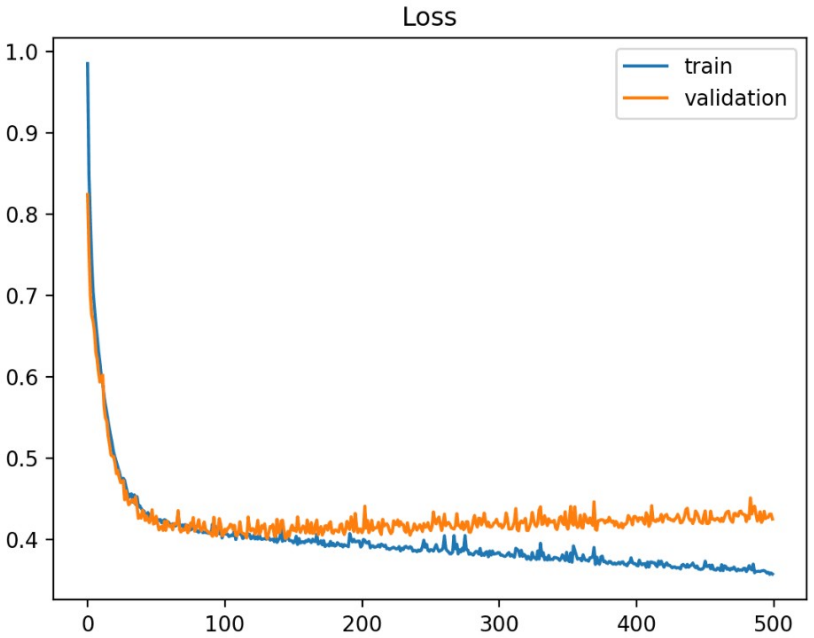
The training process can go on and possible get further improvements in learning

Overfit Model Example

The model is more complex than what is required for the problem, or the model is trained for too long.

Training and validation learning curves show overfitting if:

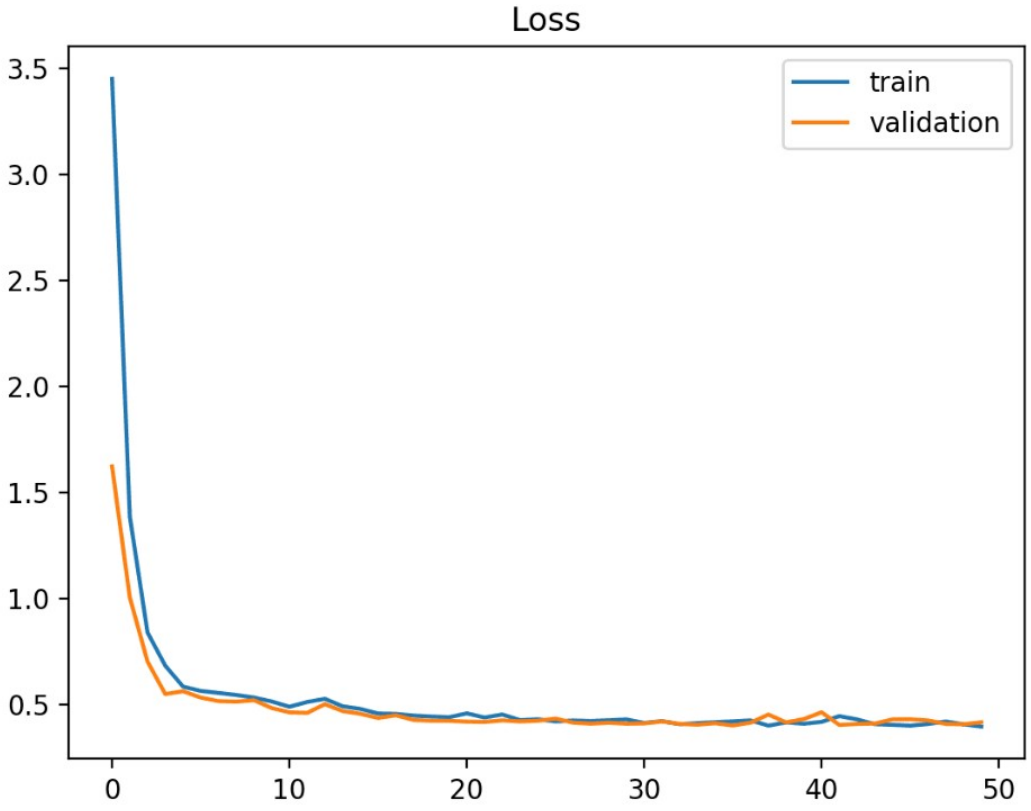
- Training loss continues to decrease with experience.
- Validation loss decreases up to a point and starts increasing again.
- The turning point in validation loss indicates the time at which training could be halted.



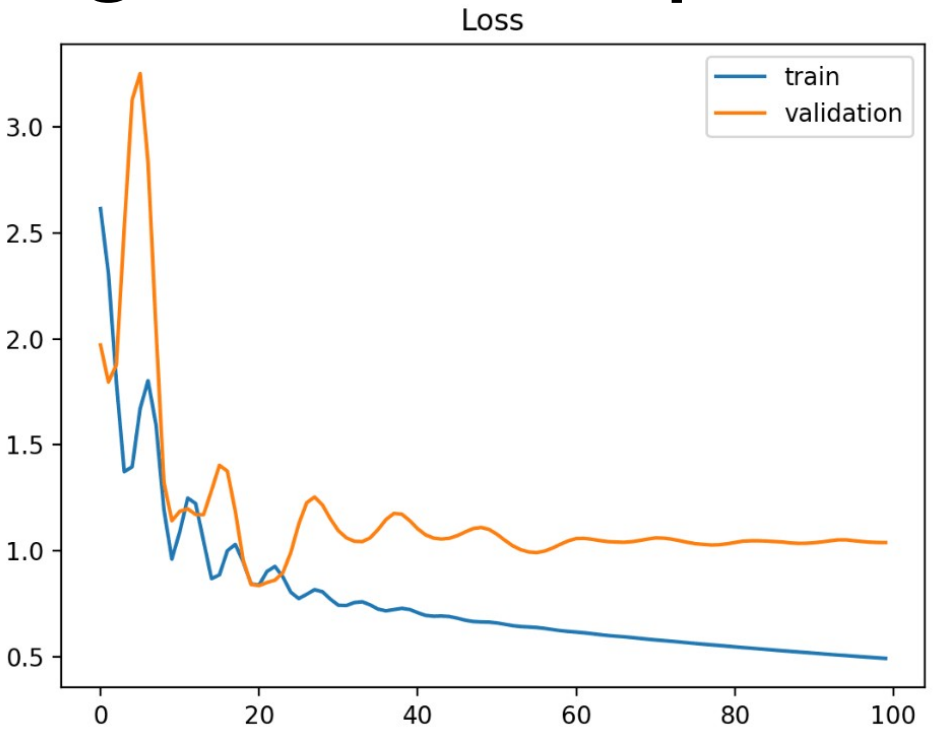
Example of Good fit

Training and validation curves decrease to a point of *stability* with a minimal gap (*generalization gap*) between the them.

The loss curve of the model is mostly lower on the training dataset than the validation dataset.



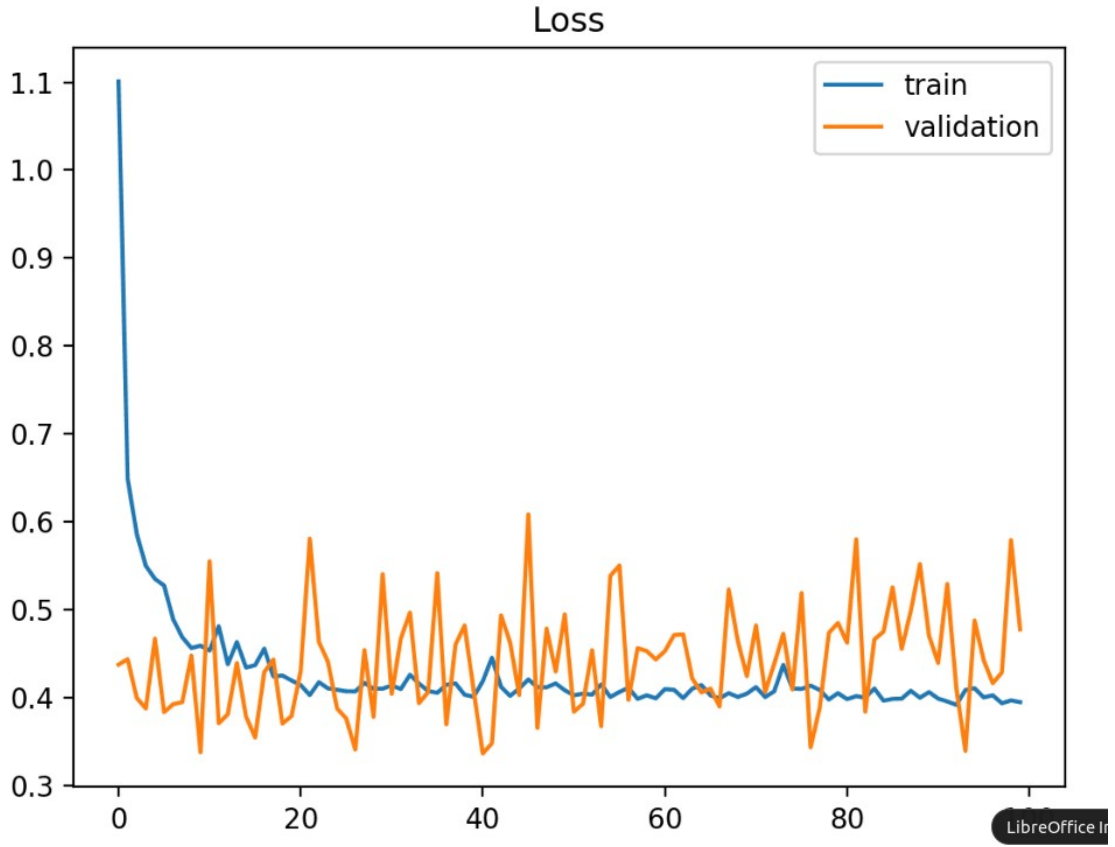
Training set not representative*



*The training data set has very few samples vs the validation set.

The curves show improvement, but there is a gap between them.

Unrepresentative* validation set



Validation set is small w.r.t. training set

Reference links

Gradient Search Visuals: https://github.com/Jaewan-Yun/optimizer-visualization

Keras Tuner: https://keras.io/keras_tuner/
Scikit Learn: https://scikit-learn.org/stable/index.html
Tensor flow: https://www.tensorflow.org/
Intro to Keras Tuner w/Tensor flow: https://www.tensorflow.org/tutorials/keras/keras_tuner

Loss curves interpretation:
https://machinelearningmastery.com/diagnose-overfitting-underfitting-lstm-models/

Multi-armed Bandit Operation