How to Establish a Baseline ML Model

ML 101 (P)

Yash Saraswat

Establish Baseline Model

Overview

- When starting to work on a Machine Learning Project, one of the most crucial step is to establish a baseline.
- Baseline helps to indicate, what might be possible.
- Usually, only after establishing a baseline level of performance, you can apply tools to efficiently improve performance of that baseline level.

Establish Baseline Model

Various Ways

- Literature Search for state-of-the-art/ Open-source
- Quick and Dirty Implementation
- Performance of Older System
- Overfit model on small dataset
- Error Analysis and Comparison against Human Level Performance

Establish Baseline Model

Literature Search

- Literature search helps find out what is possible.
- Resources for literature search can be research papers, courses, blogs, open-source projects etc.
- If you are building a machine translation system and others report a certain level of accuracy on data that's similar to your use case, then that can be your starting point.

Establish Baseline Model

Quick and Dirty Implementation

- Trying out open source implementation can be helpful in quickly creating a baseline model.
- Exploring open source implementation can give you a sense of what might be possible.
- Don’t obsess about finding latest, greatest algorithm, its better to find something reasonable that helps you get started quickly.

Reasonable algorithm with good data can over-perform a greatest algorithm with bad data.

Establish Baseline Model

Performance of Older System

- If you already have a machine learning system for your application or similar application, then performance of older system can help to establish a baseline that you can aspire to improve on.
- Transfer learning can be useful for many machine learning tasks so that you can get started quickly.
- Ex- Finetuning a T-5 model can help you quickly get started in sequence-to-sequence generation tasks.

Establish Baseline Model

Overfit on Small Dataset

- Overfitting on small dataset can be seen as sanity check for your code/algorithm.
- Before spending hours to train model on large dataset, it is better to overfit model on dataset of size 10-100.
- This can help you find potential bugs in your code much more quickly.
- If your model can't even overfit on small dataset then it is a waste of time to train model on large dataset.

Establish Baseline Model

Error Analysis and Comparison against Human Level Performance

- Error Analysis is the heart of machine learning development process.
- It is a good practice to use human level performance (HLP) as reference for doing error analysis.
- Error analysis can be done creating a table in spreadsheet or using MLOPs tools.

Error analysis can help you find the direction in which you should take next step in order to improve your model's performance.

Establish Baseline Model

Error Analysis for Speech Recognition

Type	Accuracy	HLP	Gap to HLP	% of data	Scope for Improvement
Clear Speech	94%	97%	3%	70%	2.1%
Background Noise	89%	94%	5%	4%	0.2%
Human Noise	87%	90%	3%	20%	0.6%
Low Bandwidth	71%	71%	0%	6%	0%

From the Gap-to-HLP column, it seems that Background noise category has most scope for improvement but after analysing the % of data, improving on Clear speech and Human noise seems more useful.