tds **Published in Towards Data Science**

Sophia Yang   (Follow)

Aug 24, 2022 · 7 min read · ✨ · ▶ Listen

🔖 Save   🐦   ⓕ   in   🔗

Machine Learning Model Drift

Types, causes, detections, mitigations, and tools

Photo by Robin Pierre on Unsplash

In machine learning, model drift means that the machine learning model becomes less and less accurate due to the changes in the statistical properties of the input features, target variable, or relationships among variables. The data on which the machine learning model is trained is called training data or source data. The data on which the model is trying to predict is called serving data or target data. The training/source data distributions might be different from the serving/target data distributions. In this article, we are going to walk through the types of model drift, causes of model drift, how to detect model drift, how to mitigate model drift, and finally the tools we can use to monitor model drift.

Setting

We are predicting a target variable y given a set of input features X. For example, in a house price prediction model, X could be a set of features of houses (e.g., number of bedrooms, size, location), and y could be the house price. One pair of (X, y) indicates one house record in the dataset. $p(X)$ and $p(y)$ are the probabilities of observing house features X and house price y respectively, also known as the marginal probability or prior probability. $p(y|X)$ is the conditional distribution of house prices given house features.

Types of model drift

Model drift contains two key categories: concept drift and data drift. Data drift further includes covariates/feature drift and label drift.

Covariate/feature drift

Covariate drift or feature drift happens when $p(X)$ changes but $p(y|x)$ remains the same. The marginal distribution of the input house features changes, but the conditional distribution of house prices given house features stays the same.

Let's take one house feature — size. Imagine that your model was trained before Covid and at the time, there are more larger-size houses in the market. And during Covid, people all want to move to a larger house and there are fewer and fewer larger-size houses in the market.
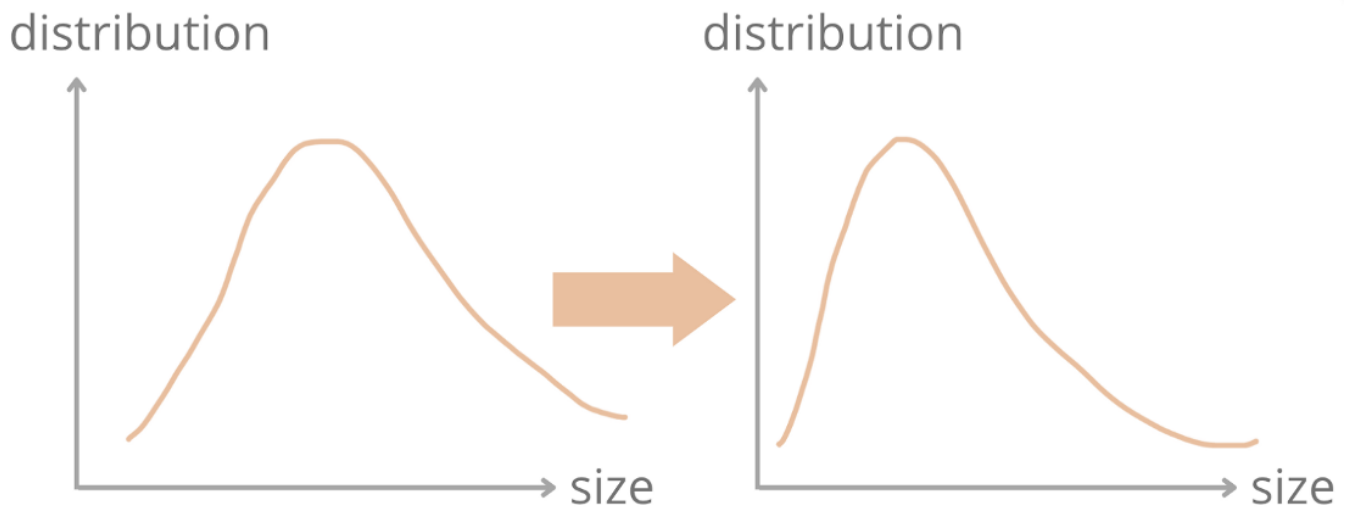
Figure 1. Covariate drift (image made by author)

Label drift

Label drift happens when p(y) changes but p(x|y) remains the same.

In the house price prediction example, the house price distribution p(y) could change after when the model was trained previously. For example, the house price has significantly increased during the pandemic, resulting in the house price distribution shifting towards a higher value.

Figure 2. Label drift (image made by author)

Concept drift

Concept drift happens when p(y|X) changes but p(X) remains the same.

In the house price prediction example, the conditional probability of housing price given house features p(y|X) could change. Let's reconsider the previous example.

Imagine that the distribution of the house sizes does not change. Because people prefer larger houses now, larger houses become more expensive. The conditional probability of housing price given house sizes could change, especially for larger houses.

Figure 3. Concept drift (image made by author)

Causes of model drift

There could be lots of causes for ML model drift. Here are a few:

One of the main causes of data drift is **sampling mismatch**. We often use some sampling strategies to find the source data to train our model. The sampling strategy could be **biased and not representative** of the entire population. For example, if we would like to develop a model to predict the housing price for the general Boston area, we can't use the training data that only include houses listed around the Back Bay area.

Sometimes people apply their well-trained model to **a new product/market/context** and find their model does not work well. For example, the pandemic changes the macro economy of the world. We are in a very different market and context compared to before the pandemic times. This is when the distribution of the features, labels, and the joint distribution of the features and labels can all change. The best solution is to retrain the models.

Anomalies can appear both in the training data and also in the target data. Anomalies can change the distributions of our data. Sometimes our models handle

anomalies automatically. Other times, we need to employ some anomaly detection techniques and filter out (e.g., winsorize) the anomalies.

Seasonal effects can also lead to a model drift. The relationship between multiple variables, for example, house size and house price, might change depending on the month or the season. It's often recommended to take into account seasonal effects in the model or in some cases, model different scenarios in different models.

Data quality issues can happen more than you would expect. There are a variety of data quality issues such as incorrect input data, incorrect data processing steps, and duplications. Sometimes the changes in data or data processing pipelines are not quality concerns but are expected and a business requirement. Such data change can also contribute to model drift if not monitored closely. Setting up guardrails to validate data and ensure data quality and consistency is a prerequisite for your ML models.

Detecting drift

Monitor model performance

The most straightforward approach to detecting drift is through monitoring model performance metrics. The most common model performance metrics include confusion matrix, accuracy, recall, F1 score, and ROC-AUC. Depending on your model usage, there might be other model behavior measure metrics that are important as well.

Along the same line, some research focuses on the error rate and uses the error rate-based drift detection method. For example, the Drift Detection Method (DDM) algorithm can be used to detect any significant increase in error rates.

Monitor descriptive statistics

We can describe our datasets statistically with measurements like min, max, median, mean, uniqueness, correlation, and others. We can include all these descriptive statistics in our model monitoring dashboard to visualize how these statistics change over time.

Monitor distribution changes

When you see changes in the descriptive statistics, how do you know there are actual changes in the distributions and how do you know those changes are

significant and meaningful? There are several statistical tests or hypothesis tests people use to detect the distribution changes statistically such as Population Stability Index, Kullback-Leibler (KL) divergence, Jensen-Shannon, Kolmogorov-Smirnov test, Cramér-von Mises, Fisher's Exact Test, Maximum Mean Discrepancy (MMD), and Least-Squares Density Difference.

For example, **Population Stability Index** (PSI) measures "how much a population has shifted over time or between two different samples of a population in a single number". **Kolmogorov-Smirnov test** (or KS test) "quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution, or between the empirical distribution functions of two samples".

Furthermore, there are also supervised and unsupervised ML models designed for detecting model drift.

wrong with our models. In either situation, when model drift happens, we first need to find the cause of the drift. What are the causes of the model drift? Different root causes should have different mitigation strategies. Please refer back to the "Causes of model drift" for details.

Assuming our data is correct and validated, a common approach to mitigate model drift is to retrain the model with new data. You may wonder, why don't we just retrain models all the time so that we don't need to care about model drift? That's a great question. In fact, a lot of companies do model retraining as much as they can. However, sometimes model retraining and deployment are not that straightforward. For example, your new data might not be labeled. Your newly trained model might not be better, especially when you have not done the drift analyses and are not sure where it went wrong in the first place.

Tools

There are a lot of ML monitoring tools and model drift monitoring tools out there. Here are a few popular tools:

whylogs is an "open source library for logging any kind of data. With whylogs, users are able to generate summaries of their datasets (called whylogs profiles) which they can use to track changes in their dataset, create data constraints to know whether their data looks the way it should, and quickly visualize key summary statistics about their datasets."

Evidently is an "open-source framework to evaluate, test and monitor ML models in production."

Alibi Detect is an "open source Python library focused on outlier, adversarial and drift detection."

All the major cloud providers and MLOps platforms also have their own model performance and model drift monitoring tools if you use those platforms. For example, Amazon SageMaker Model Monitor "detects data skew by comparing real-world data to a baseline dataset such as a training dataset or an evaluation dataset."

Conclusion

In your machine learning journey, designing a machine learning model is just the start. The real challenge follows once you have crafted your perfect machine learning model and deployed it into production. The model that works for the current data might not work for the future. A big component of continuous monitoring is model drift. Monitoring and mitigating model drift help to keep your machine learning model happy and successful!

References:

- A survey on concept drift adaptation. *ACM computing surveys (CSUR)*

- Quionero-Candela, Joaquin, et al. *Dataset shift in machine learning*. The MIT Press, 2009.

- Design Machine Learning System. Chip Huyen. 2022.

- https://www.fiddler.ai/blog/drift-in-machine-learning-how-to-identify-issues-before-you-have-a-problem

- https://mwburke.github.io/data science/2018/04/29/population-stability-index.html

. . .

By Sophia Yang on August 23, 2022.

Sophia Yang is a Senior Data Scientist at Anaconda. Connect with me on <u>LinkedIn</u>, <u>Twitter</u>, and <u>YouTube</u> and join the DS/ML <u>Book Club</u> ❤️

Machine Learning   Data Science   Data Drift

Sign up for The Variable

By Towards Data Science

Every Thursday, the Variable delivers the very best of Towards Data Science: from hands-on tutorials and cutting-edge research to original features you don't want to miss. <u>Take a look.</u>

By signing up, you will create a Medium account if you don't already have one. Review our <u>Privacy Policy</u> for more information about our privacy practices.

[✉️⁺ Get this newsletter]

About   Help   Terms   Privacy

Get the Medium app

[Download on the App Store]   [GET IT ON Google Play]