

Join us on 2/15 for AI Explained: The Ins and Outs of Foundation Models with Alex Ratner, Co-founder and CEO of Snorkel AI. Register

[Product](#) ▼[Solutions](#) ▼[Pricing](#) ▼[Request demo](#)[Resources](#) ▼[Company](#) ▼[← Back to blog home](#)[ML MODEL MONITORING](#)[MLOPS](#)

# Drift in Machine Learning: How to Identify Issues Before You Have a Problem

Inaccurate models can be costly for businesses. Whether a model is responsible for predicting fraud, approving loans, or targeting ads, small changes in model accuracy can result in big impacts to your bottom line. Over time, even highly accurate models are prone to decay as the incoming data shifts away from the original training set. This phenomenon is called **model drift**.

Here at Fiddler we want to empower people with the best tools to monitor their models and maintain the highest degree of accuracy. Let's dig into what causes model drift and how to remedy it. You can also hear about model drift directly from me in this [video](#).

ML Drift: Identifying Issu...



PUBLISHED JANUARY 31,

Amy Hodler

SHARE



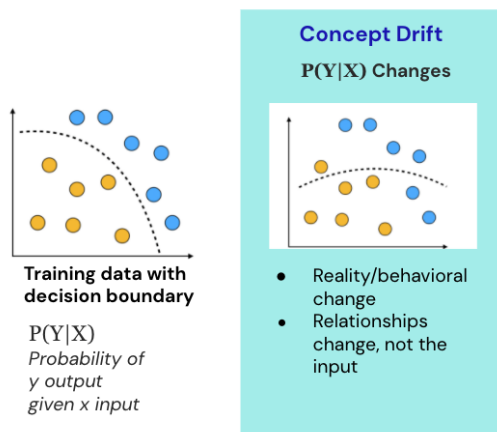
Subscribe  
to Fiddler  
newsletter

Business Email

Subscribe

# Types of model drift

When we talk about model drift, there are three categories of changes that can occur. Keep in mind these categories are not mutually exclusive. We'll walk through each category and describe it using examples from a model that is designed to assess loan applications.



Concept drift occurs when actual behavior changes.

Plot source: [Don't let your model's quality drift away](#) by Michał Oleszak

## Concept drift

Concept drift indicates there's been a change in the underlying relationships between features and outcomes: the probability of  $Y$  output given  $X$  input or  $P(Y|X)$ . In the context of our loan application example, concept drift would occur if there was a macro-economic shift that made applicants with the same feature values (e.g. income, credit score, age) more or less risky to loan money to. The plot shows data with two labels – orange and blue (potentially loan approvals and non-approvals). When concept drift occurs in the second image, we observe a new decision boundary between orange and blue data as compared to our training set.

## Data drift

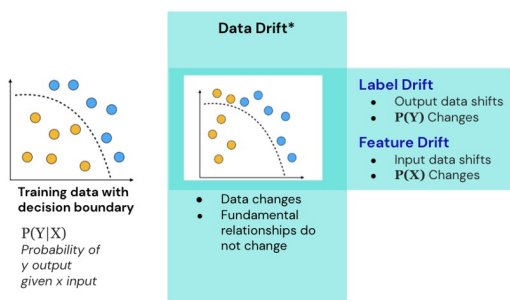
Data drift refers simply to changes we observe in the model's data distribution. These changes may or may not correspond to a new relationship between the model's features and outcomes. Data drift can be further categorized as feature drift or label drift.

# Feature drift

Feature drift occurs when there are changes in the distribution of a model's inputs or  $P(X)$ . For example, over a specific time frame, our loan application model might receive more data points from applicants in a particular geographic region. In the image above, we observe more orange data points towards the smaller end of the x-axis as compared to the training set.

# Label drift

Label drift indicates there's been a change in a model's output distribution or  $P(Y)$ . If we see a higher ratio of approval predictions to non-approval predictions, this would be an example of label drift. On our plot, we see some of the orange data points higher on the y-axis than the training data that are now on the "wrong side" of the decision line.

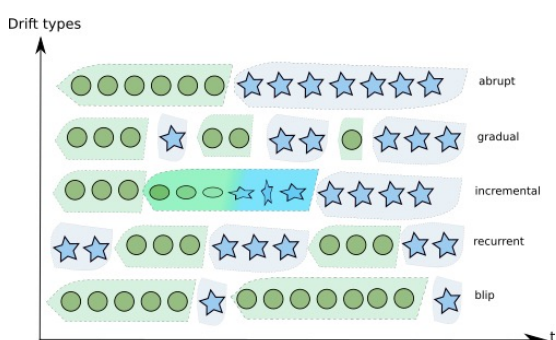


Data drift occurs when the data distribution changes.

Plot source: [Don't let your model's quality drift away](#) by Michał Oleszak

Feature drift and label drift are inherently related to concept drift via Bayes' theorem. However, it's possible to observe data drift without observing concept drift if the shifts balance out in the equation. In this case, it is still important to identify and monitor data drift, because it could be a signal of future performance issues.

# How to detect model drift



Our experience of drift differs.

Image source: KD nuggets [The ravages of concept drift](#)

Model drift can occur on different cadences. Some models shift abruptly — for example, the COVID-19 pandemic caused abrupt changes in consumer behavior and buying patterns. Other models might have gradual drift or even seasonal/cyclic drift.

Regardless of how the drift occurs, it's critical to identify these shifts quickly to maintain model accuracy and reduce business impact.

If you have labeled data, model drift can be identified with performance monitoring and supervised learning methods. We recommend starting with standard metrics like accuracy, precision, False Positive Rate, and Area Under the Curve (AUC). You may also choose to apply your own custom supervised methods to run a more sophisticated analysis. Learn more about these methods in [this](#) review article.

If you have unlabelled data, the first analysis you should run is some sort of assessment of your data's distribution. Your training dataset was a sample from a particular moment in time, so it's critical to compare the distribution of the training set with the new data to understand what shift has occurred. There are a variety of distance metrics and nonparametric tests that can be used to measure this, including the Kullback-Leibler divergence, Jensen-Shannon divergence, and Kolmogorov-Smirnov test. These each have slightly different assumptions and properties, so pick the one that's most appropriate for your dataset and model. If you want to develop your own unsupervised learning model to assess drift on unlabelled data, there are also a number of models you can use. Read more about unsupervised methods for detecting drift [here](#).

## Determining the root cause

You've identified that model drift is occurring, but how do you get to the root cause?

Drift can be caused by changes in the world, changes in the usage of your product, or data integrity issues — e.g. bugs and degraded application performance. Data integrity issues can occur at any stage of a product's pipeline. For example, a bug in the frontend might permit a user to input data in an incorrect format and skew your results. Alternatively, a bug in the

backend might affect how that data gets transformed or loaded into your model. If your application or data pipeline is degraded, that could skew or reduce your dataset.

If you notice drift, a good place to start is to check for data integrity issues with your engineering team. Has there been a change in your product or an API? Is your app or data pipeline in a degraded state?

The next step is to dive deeper into your model analytics to pinpoint when the change happened and what type of drift is occurring. Using the statistical tools we mentioned in the previous section, work with the data scientists and domain experts on your team to understand the shifts you've observed. Model explainability measures can be very useful at this stage for generating hypotheses.

Depending on the root cause, resolving a feature drift or label drift issue might involve fixing a bug, updating a pipeline, or simply refreshing your data. If you determine that context drift has occurred, it's time to retrain your model.




## Staying on top of drift

We've given a brief overview of the different types of model drift and how to identify them. All models are subject to decay over time, which is why it's critical to be aware of drift and have appropriate tools to manage it.

At Fiddler, we believe in Responsible AI, and maintaining model accuracy is core to our philosophy. Fiddler offers a centralized management platform that continuously monitors your AI and produces real-time alerts when there are signs of drift. We also provide a suite of tools in-app to assess model performance and generate explainability measures. [Request a demo](#) to learn more about how Fiddler can help protect your company's AI from drift.

# Ready to try model performance management?

Request Demo

PRODUCT	SOLUTIONS	RESOURCES	COMPANY	CONNECT
Model Performance Management	Customer Experience	Blog	About Us	Events
		Docs	Careers <small>WE'RE HIRING!</small>	Slack
Model Monitoring	Lending and Trading	Resource Library	Newsroom	Community
	Lifetime Value		Partners	Subscribe to Newsletter
Explainable AI		Guides	Amazon SageMaker + Fiddler	Contact Sales
Analytics	Risk and Compliance	Research		Support Center
Fairness		Videos		  
MLOps		What is Model Monitoring?		
Build vs Buy				
NLP and CV Model Monitoring				
Security				