Tecton Feature Store Helps Tide Deploy Models 2x Faster with 3x More Features

Tide is a UK-based mobile-first banking platform for small- and medium-sized enterprises (SMEs), offering business bank accounts with quick on-boarding, low fees, and a range of innovative solutions for managing banking and admin. Founded in 2015 to simplify business banking, Tide has grown to serve over 300,000 members in 5 years, with over 600 employees and offices in London, Sofia (Bulgaria), and Hyderabad (India).

Tide thrives on making data-driven decisions to help their customers save time and money. To support this objective, a dedicated team of data scientists and engineers develop and productionize Machine Learning (ML) models that power automated experiences across Tide's customer ecosystem. Tide aspires to automate as much of their business decision-making as possible with operational ML.

Before engaging with Tecton, Tide had started working on ML-based fraud detection to identify fraudulent transactions and risk assessments to automate new account approvals. They also planned to launch a production model to match members' transactions with their outstanding invoices. For each of these use cases, Tide found it critical to be able to make predictions online in real-time: to detect fraud as quickly as possible, minimize friction in customer onboarding, and resolve invoices as close to the point of transaction time as possible. This meant that Tide faced the challenge of serving features online based on streaming data—their existing models used batch data and offline features only.

Company:

tide

Industry:

Business banking

Challenges:

Due to the complexity of building data pipelines, Tide was struggling to scale its operational ML efforts:

» Models took 2–4 months to deploy due to handoffs between data scientists and engineering

» Using streaming data sources required building and maintaining bespoke data pipelines

» Batch data and streaming data were difficult to combine into one system

» Completing and maintaining an internal feature store required 6–9 months and 3 new FTEs

Solution:

Tide implemented the Tecton enterprise feature store to build and deploy ML features as a core component of Tide's ML stack. Tecton is now used for multiple production use cases and will support additional use cases in the future:

» In production: Risk detection for customer onboarding, invoice matching for transactions, invoice default predictions, balance predictions

To address real-time use cases and scale their operational ML efforts across the board, Tide began work on an internal feature store. They were able to start serving some of their features online, but encountered hurdles incorporating both batch and streaming data into their feature store. Tide estimated that completing a basic feature store to meet their requirements would take 6–9 additional months, and to maintain it, they would need to hire 3 full-time employees (FTEs).

Considering the costs of building their own solution, Tide decided to evaluate Tecton's enterprise feature store. With support from Tecton's engineering team, Tide was able to set up Tecton's fully integrated feature store platform in production in just 6 weeks. After a successful proof-of-concept, Tide has been able to cut the time it takes to deploy a model from 2–4 months to just 1 month, deploy 2x more models than they had previously, and release intelligent new products for tens of thousands of their business customers.

ML Use Case Priorities

Tide's data science and engineering team is focused on using operational ML to make immediate, real-time decisions for their customers, specifically in the area of risk evaluation. One important use case is reviewing new customers during onboarding to determine whether to automatically approve the account. This should happen with a benchmark of less than 200 milliseconds of latency so that the customer does not experience delay during the sign-up process.

Tide has also built models to detect other types of risk, such as fraudulent transactions across the hundreds of thousands of transactions that occur daily on their platform. These models have high feature freshness requirements; they require streaming data to accurately assess the risk of recent transactions. This data is also critical for another ML use case, invoice matching, where Tide will intelligently pair incoming transactions with open invoices to save their customers time—resolving invoices can easily take a small business 10–20 hours per month.

» In development: Fraud detection on transactions, insolvency predictions

» On the roadmap: Receipt information extraction, industry classification, recommendation engine, transaction classification

Results:

Tide is already seeing significant value from implementing Tecton's enterprise feature store:

» Reduced time to deploy new models by more than 50%, from 2–4 months down to just 1 month

» Increased the number of features per model by 3x through enabling sharing and reuse of features across models

» Empowered data scientists to build production-ready features and deploy them to production instantly

» Improved the customer experience for 42,000 companies by launching invoice matching with 97% model accuracy

» Recovered $600K annually by drastically reducing the need for manual review of new accounts

» Saved 3 FTEs to maintain and 6–9 months investment to build an internal feature store

» Introduced management of "features as code" and brought DevOps-like practices to feature engineering

Challenges Getting ML to Production

Tide is a large-scale business:

- Over 300,000 customers
- Millions of events generated every day
- Millions of feature-key combinations per model
- 1M individual predictions generated every day

With data of this magnitude, getting the first models to production was a challenge. Without a feature store in place to manage the lifecycle of ML features, Tide's team was struggling with:

> **Lack of collaboration tools spanning data science and engineering:** Data scientists and engineers spent substantial time on manual coordination. Lacking shared feature definitions spanning production and development environments, the team was not able to reuse features across models. They had to reproduce work and/or use fewer features, potentially at the cost of model accuracy.
>
> **Lead times of 2–4 months to productionize models:** One of Tide's greatest challenges was long cycle times for productionizing models. Due to a lack of collaborative tools, data scientists at Tide didn't have a way to build production-ready features, and when they wanted to add more data to improve a model, they had to coordinate with engineering. Handoffs significantly increased the time it took to get a model to production.
>
> **Challenge of incorporating streaming data:** Tide's initial models used batch data sources only, but new customer use cases required real-time predictions and streaming data. Tide needed to integrate its predictive models with Kafka and develop a way to handle issues around using real-time sources, such as the potential for data leakage in training. They would also have to serve features online at a low latency (p99 < 200ms) and with a load of 5–20 QPS.

To solve these issues, Tide started to build an internal feature store. While they were able to use streaming data to generate real-time predictions, they quickly encountered challenges, such as ingesting both batch and streaming data sources into their online feature store. Tide estimated that there would be significant costs for continuing on this path: spending 6–9 months to complete their internal feature store, and hiring at least 3 new FTEs to maintain it.

Solution

Tecton was founded by the creators of Uber Michelangelo and provides an enterprise-ready feature store. When Tide was evaluating commercial solutions, they became interested in Tecton because it covered their existing needs and offered functionality they weren't likely to build in-house, like data monitoring, that would be valuable as Tide continued to scale operational ML.

Tide and Tecton ran a proof-of-concept (POC) project. Over the course of 6 weeks, the teams worked together to implement the Tecton enterprise feature store and test several key use cases on Tecton, including risk detection for customer onboarding and invoice matching for transactions. After the successful POC, Tide is now using Tecton for the majority of its models in production, and has decided to deploy the Tecton feature store as a core component of Tide's ML stack for future use cases.

Results

> *"Tecton allowed us to significantly speed up our model productionization times, enabling us to get faster feedback on our ML applications and in turn build products that deliver substantial value to our customers."*
>
> *Hendrik Brackmann, Director of Data Science and Analytics at Tide*

As a result of deploying Tecton's enterprise feature store, Tide has almost doubled the number of models in production. They launched a production model for invoice matching with 97% accuracy, serving ~42,000 companies and producing tens of thousands of matches per day. Tide has also launched risk detection for customer onboarding, with an estimated savings of $600K by drastically reducing the need for manual review of new accounts. Additional results include:

Total time to deploy a model reduced by more than 50%: In the past, it would take the data science and engineering team 2-4 months to get a model into production. With Tecton, that time has been reduced to around 1 month.

Enabled sharing of features across projects: After just a short period of time, Tide is already seeing features reused across models due to the benefit of a centralized feature repository and standardized feature definitions. This means that models are able to use on average 3x as many features as they were previously, increasing model accuracy.

Instant provisioning of features to production: Tecton automated Tide's feature pipelines from definition, through transformation and storage, and into serving, enabling Tide's data scientists to push features to production instantly. Using Tecton, data scientists can build production-ready features quickly and are empowered to own their work end-to-end.

Built-in capabilities for scaling operational ML: As Tide continues to grow its operational ML efforts, they are interested in using Tecton's data monitoring and built-in support for higher data loads. These are capabilities that Tide was unlikely to build internally, and would have been costly to support long-term.

Mitigated data leakage: Tecton's simple time-travel feature backdating capabilities enable Tide's data scientists and engineers to easily and reliably avoid data leakage in model training and consistently generate time-accurate data sets.

Avoided hiring 3 FTEs and accelerated feature store deployment by 6–9 months: Tide had estimated that their own internal feature store would require 6–9 months to complete and 3 new FTEs to maintain. With Tecton, they were able to get a full feature store implementation deployed in 6 weeks with no need to increase their hiring budget.

Tide's vision is to use operational ML to create high-value customer experiences and automate internal business processes. To achieve this, Tide needed a full stack for operational ML to ingest more data sources, help data scientists and engineers collaborate, and reduce the time to deploy new models. With Tecton's enterprise feature store as the data layer for their ML stack, Tide now has the infrastructure in place to deliver on their product vision.