

Forecasting a \$1B+ consumer-health portfolio at >97% accuracy

Previsión de demanda para un portafolio de más de 1.000 M\$ con más del 97 % de precisión

Role: Senior Data Science & Analytics Lead · Pfizer Consumer Healthcare, Madison NJ · 2016 – 2021 (through the Pfizer CHC / GSK CHC merger that formed Haleon)

CONTEXT

Consumption and demand forecasts for a \$1B+ OTC portfolio were produced manually, category by category, with little use of external drivers. Supply planning, finance and brand teams each held their own numbers. Accuracy degraded around seasonality (cough/cold, allergy) and around promotions.

Alongside forecasting, the business had no scalable way to read what consumers were saying about products online.

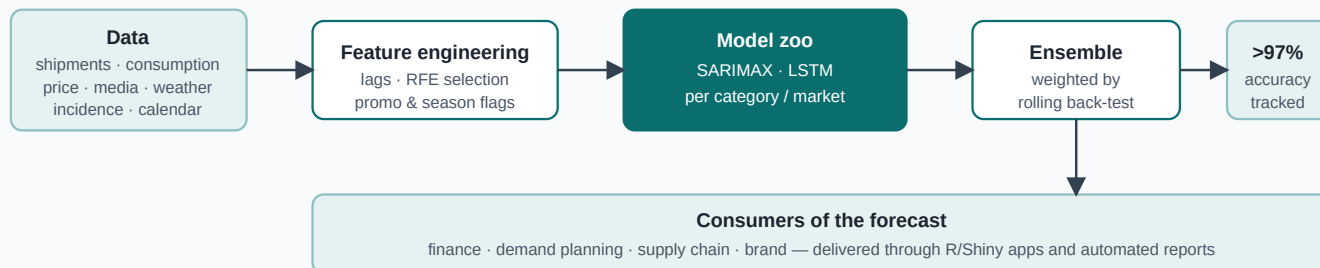
WHAT I BUILT

- An end-to-end ML pipeline: automated ingestion of shipments, consumption, pricing, media, weather and disease-incidence data; feature engineering with recursive feature elimination; a model zoo of SARIMAX and LSTM models; ensembling and back-testing; scheduled production runs.
- Full-stack analytics products with a Python back end and R/Shiny front end: portfolio optimization, product-assortment tools and an NLP review-analytics app.
- Social-listening trend detection with NLP on unstructured data, feeding brand strategy and the innovation pipeline.
- Browser automation (RSelenium/PhantomJS) to remove manual third-party data pulls.

OUTCOME

- **>97% forecast accuracy** sustained across the portfolio; one shared number for finance, demand planning and supply chain.
- Forecasts, MMM and pricing analytics embedded in commercial decision-making; I acted as the analytics partner to finance and supply chain leadership.
- Managed the data science team and defined KPIs and measurement frameworks; applied causal methods to quantify campaign effectiveness.
- Recognized with the Pfizer Research Award.

FORECASTING PIPELINE



Companion products: NLP review analytics · social-listening trend detection · portfolio optimization · assortment tools · Bayesian MMM & digital attribution (see case study 2)

>97%

forecast accuracy, SARIMAX + LSTM ensembles

\$1B+

portfolio covered by the forecasting pipeline

\$100M+

media budgets informed by Bayesian MMM

Award

Pfizer Research Award recipient

STACK

Python

R / Shiny

SQL

SARIMAX

LSTM

PyMC3

Stan

NLP

RSelenium

Power BI

Sanitized: describes methods and my role only — no brand-level volumes, prices or internal forecasts are disclosed. Pfizer is a trademark of Pfizer Inc., referenced here solely to identify where the work was carried out.