Game Sales Forecasting Engine

Brief Proposal Document

A fully reproducible version of this forecasting engine, including code, feature pipeline, and Power BI dashboards is available in the accompanying GitHub repository link. (https://github.com/georgejp144/game-sales-forecasting-engine)

1. Introduction

The global games market is increasingly hit-driven, competitive, and volatile. Yet the forecasting practices used by most publishers still rely on:

- Past-title analogues
- Franchise averages
- Simple ratios and static decay models

These legacy approaches miss the complex, multifactor interactions that shape real-world sales trajectories across launch, mid-tail, and long-tail phases.

This project introduces a Predictive Game Sales Forecasting Engine, a modern multi-layer forecasting system inspired by methods used in quantitative finance, time-series modelling, and cross-sectional machine learning.

It is built to deliver:

- Accurate weekly forecasts over a full 52-week lifecycle
- Cold-start capability (AAA / AA / Indie / New IP)
- Scenario ranges (P10 / P50 / P90)
- Explainability through an integrated Power BI dashboard

The engine replaces heuristic forecasting with a structured, transparent, risk-aware model suitable for commercial planning, marketing optimisation, and release strategy.

2. Business Problem

Modern publishing teams must plan marketing, pricing, inventory, promotions, and DLC schedules in an environment where sales behaviour is volatile and difficult to anticipate. Two fundamental problems drive this uncertainty:

2.1 Cold Starts & Non-Linear Sales Behaviour Make Forecasting Unstable

New releases, especially new Ips, launch with no sales history, forcing reliance on genre averages or franchise analogues that fail to capture real differences in:

- Marketing intensity
- Franchise strength
- Seasonal timing
- Platform behaviour
- Competitive release windows

Once live, real-world game sales rarely follow smooth decay. They are driven by:

- DLC drops
- Discount campaigns
- Review/viral sentiment
- Platform featuring
- Competitor releases

Traditional curve-fitting or ratio-based methods cannot adapt to this event-driven volatility, making forecasts inconsistent, slow to update, or structurally unrealistic.

How the engine solves this:

The behavioural prior curve provides a realistic structural baseline for each dev-type, while cross-sectional transfer learning (XSTL) allows new titles to borrow behaviour from similar historical games. Residual XGBoost then learns deviations from the baseline caused by DLC, discounts, marketing, and competition, while promo-safe smoothing prevents unrealistic spikes, giving stable, realistic forecasts across both cold-start and live phases.

2.2 Commercial Teams Need Reliable, Explainable Forecasts

Marketing, finance, production, and distribution teams must make decisions that depend on:

- High/medium/low scenario planning
- Confidence intervals
- Weekly reforecasting
- Understanding what drives changes
- Knowing when to trust or distrust the forecast

Legacy models provide single values without:

- Uncertainty
- Transparency
- Reliability scoring
- Driver attribution
- Diagnostics or regime awareness

This leaves teams with forecasts that are hard to justify, audit, or use in commercial planning.

How the engine solves this:

The system outputs P10/P50/P90 ranges, a Forecast Uncertainty Index, and weekly updates. The reliability framework (regime confidence, drift score, and XSTL similarity) quantifies trust in the prediction. Clear baseline vs blended curves and marketing uplift metrics make the forecast fully explainable and commercially actionable.

3. Model Architecture

The engine uses a multi-layer architecture where each stage addresses a specific forecasting weakness.

3.1 Behavioural Prior Curve (Real Space)

A structural lifecycle model that provides the baseline expectation for any title.

Equation from proposal:

$$Sales(t) = Peak \times e^{-kt} \times Seasonal(t)$$

Extended with parameters for:

- Dev-type decay patterns (AAA / AA / Indie)
- Seasonality uplift
- DLC uplift + 2-week echo
- Discount uplift + 1-week echo
- Indie late-tail flattening (post Week 20)

This layer ensures that every forecast starts from a realistic, commercially recognisable shape, especially critical for cold starts.

3.2 Residual XGBoost (Log Space)

Instead of predicting sales directly, the model predicts log-residuals:

$$\text{Residual} = \log\left(1 + \text{Sales}_{\text{actual}}\right) - \log\left(1 + \text{Prior}\right)$$

Converted back:

$$\text{Sales}_{\text{pred}} = \exp(\log(\text{prior}) + \text{residual}) - 1$$

XGBoost learns deviations from expected behaviour including:

- Marketing effects
- Competition pressure
- DLC/discount timing
- Seasonal interactions
- Cross-title patterns

Residual learning stabilises training, controls variance, and massively reduces overfitting.

3.3 Cross-Sectional Transfer Learning (XSTL)

The new title is compared against 200 synthetic historical titles using:

- Weekly cross-title statistics
- Dev-type and franchise patterns
- Mahalanobis distance in feature space

This produces:

- XSTL similarity score
- Cross-sectional feature set

XSTL is essential for cold-start accuracy, allowing the model to borrow behaviour from similar historical titles.

3.4 Reliability Scoring (0–1)

Three internal diagnostics:

- Regime Confidence
- Feature Drift Score
- XSTL Similarity

Combined:

$$\text{Reliability} = \left(\mathrm{R}_{\text{regime}} \times \mathrm{R}_{\text{drift}} \times \mathrm{R}_{\text{xstl}}\right)^{1/3}$$

Used to dynamically weight:

- Prior Curve
- Promo-safe XGB Output

High reliability → XGB weight up to ~0.90
Low reliability → Prior dominates (min XGB weight = 0.30)

3.5 Forecast Outputs

Each weekly forecast includes:

- Final blended prediction
- P10 / P50 / P90 uncertainty bands
- Forecast Uncertainty Index (P90–P10)
- Reliability score
- Baseline no-marketing curve
- Marketing uplift
- Full metadata for Power BI

These outputs allow marketing, finance, and commercial teams to evaluate both the prediction and its confidence.

4. Example Forecast: NeonRift (AAA) Demo

To demonstrate end-to-end system behaviour, the proposal walks through the AAA title NeonRift across its full 52-week lifecycle.

4.1 Prior Curve (Structural Baseline)

NeonRift prior parameters:

- Peak: 1.45M
- k-decay: 0.087
- Franchise strength: 0.888
- Dev-type: AAA
- Launch seasonality: 1.448

Early week priors:

Week	Prior
0	1,450,000
1	1,328,617
2	1,217,395
3	1,115,484

This reflects a strong, front-loaded AAA curve with realistic early decay.

4.2 Residual XGB Predictions

Residuals learned from marketing, competition, DLC flags and seasonality:

Week	Residual	Raw XGB
0	+115,476	1,565,476
1	+114,254	1,1442,871
2	+94,671	1,312,066
3	+106,012	1,221,496

These adjustments capture marketing momentum, competition shifts, and cross-title similarity.

4.3 Reliability-Weighted Blending

The final forecast blends Prior and XGB based on Reliability.

Week 0 reliability:

- Regime: 0.2259
- Drift: 0.0215
- XSTL: 0.0638
- Composite: 0.06766
- XGB Weight: 0.30 (minimum)

Week 7 reliability:

- Composite: 0.12449
- XGB Weight: 0.32407

Across Year 1:

XGB weight ranges 0.30 → 0.59, with the prior stabilising low-confidence periods and ML dominating when reliability improves.

4.4 DLC, Discount and Promo Smoothing

NeonRift promotional events:

- DLC: Weeks 20, 32, 44
- Discounts: Weeks 26, 40

The engine applies:

- Uplift caps
- 1–2 week echo smoothing
- Prior protection

This prevents unrealistic promo spikes and produces smooth mid- and late-tail behaviour.

4.5 Final Blended Forecast

Sample of Weeks 0–9:

Week	Final Blend
0	1,484,642
1	1,362,893
2	1,245,796
3	1,147,287

A smooth, structurally consistent curve that adapts as information increases.

4.6 Uncertainty (P10–P90)

Using symmetric scaling:

- P10 = 0.85 × blended
- P90 = 1.15 × blended

Examples:

Week	P10	P90	Spread
0	1,261,946	1,707,339	445,393
1	1,158,459	1,567,327	408,868
2	1,058,927	1,432,666	373,739

4.7 Marketing Uplift

- Week-0 baseline: 965k
- Week-0 blended: 1.48M
- Uplift: ~520k units

This uplift remains interpretable across the entire lifecycle.

4.8 Lifecycle Totals

- Launch (0–4): ~6.2M
- Mid-tail (5–20): 10–14M
- Long-tail (21–52): ~17.6M total

5. Power BI Dashboard Summary (NeonRift)

Below is a summary of the NeonRift dashboard, which transforms model outputs into clear, commercially actionable insights.

Power BI dashboard images included within a set of accompanying PDF files.

Individual Dashboard Visuals

1. KPI Header

Shows:

- Dev type: AAA
- Franchise and strength
- Decay class and rate
- Platform and region
- Total Annual Sales
- Marketing Uplift
- Week 1 sales
- Sales Half-Life Week
- Sales Plateau Week
- Forecast Uncertainty
- SMAPE

Gives executives an instant overview of commercial performance.

2. Baseline vs Blended Curve (with XGB Uplift)

- Baseline curve = pure prior
- Blended curve = final prediction
- XGB Uplift shows residual corrections

This visual explains exactly how ML is adjusting the baseline.

3. Cumulative Forecast (Lifecycle Segmentation)

Displays cumulative sales with vertical markers for:

- Launch (0–4)
- Mid-Tail (5–20)
- Long-Tail (21–52)

Shows whether the title is front-loaded or tail-weighted.

4. Uncertainty Bands (P10–P90)

Shaded area shows weekly forecast volatility.

5. Promotion Impact Timeline

DLC (gold) and discount (green) bars aligned with the final curve. Useful for assessing promo effectiveness and pricing strategy.

6. Marketing ROI vs Baseline

Direct comparison of blended curve vs no-marketing baseline. Shows incremental units from marketing week-by-week.

7. Model Reliability Over Time

Plots reliability score with DLC/discount markers. Indicates when the forecast is most trustworthy.