IRB EAD Modeling: Technical Documentation and Model Report

Melina Hafelt

July 10, 2025

Contents

1 Abstract

This report outlines the methodology, results, and regulatory considerations of an **Exposure at Default (EAD)** modeling exercise under the IRB approach. The model is developed on a synthetic dataset designed to replicate real-world credit line utilization behavior prior to default.

Highlights:

- Structured feature engineering and outlier treatment
- Model development with XGBoost and performance benchmarking
- Residual diagnostics and calibration by product/segment/quintile
- Stress testing with scenario-based simulations
- IRB-aligned documentation, controls, and model governance

Note: All data used are simulated and do not reflect real customer exposures.

2 Theoretical Background

2.1 Definition of EAD

Exposure at Default (EAD) refers to the expected value of the outstanding exposure at the time of a borrower's default. For revolving credit facilities, it may exceed the current exposure due to drawdowns before default. The general formula is:

$$\text{EAD} = \text{Current Exposure} + \text{Undrawn Amount} \times \text{CCF} \tag{1}$$

Where CCF is the Credit Conversion Factor.

2.2 Probabilistic Nature of EAD

EAD estimates are **expected values**, not point predictions. Some obligors will default with higher, others with lower exposures. The model aims for accuracy at portfolio and segment levels.

2.3 Regulatory Context

Per Basel II/III and CRR Article 166, EAD models should:

- Reflect potential increase in exposure before default
- Use CCFs based on historical utilization patterns
- Ensure conservatism under adverse conditions

3 Data Preparation and Quality Review

3.1 Encoding and Feature Engineering

- **Categoricals encoded**: product type, segment
- **Utilization rate calculated**: current exposure / credit limit
- **Nulls and duplicates:** None found
- **Outliers:** None extreme; realistic distribution (1st–99th pct: 9k–71k)

3.2 Sanity Checks

- EAD $\geq$ current exposure: ✓
- Credit limit $\geq$ current exposure: ✓
- CCF $\in [0, 1]$: ✓

4 Model Development

4.1 Features Used

- credit_limit
- current_exposure
- ccf
- utilization_rate
- product_type_encoded
- segment_encoded
- months_on_book
- num_transactions_last_month

4.2 Model and Metrics

- Model: **XGBoost Regressor**
- Train-test split: 80/20
- $\mathbf{R}^2$: 0.87
- **RMSE**: 5,118 kr

4.3 Residual Analysis

- Residuals are centered around 0
- No evidence of systematic under- or overestimation
- Slight variance increase at higher predicted values

4.4 Feature Importance

Top drivers based on XGBoost gain:

- product_type_encoded
- current_exposure
- credit_limit
- utilization_rate

5 Calibration and Diagnostics

5.1 Prediction Error by Group

Errors grouped by product type, segment, and EAD quintile. No segment showed large structural bias. Calibration plots confirm prediction quality is stable across portfolio.

6 Stress Testing

6.1 Scenarios Simulated

- **CCF +20%**: Simulate aggressive drawdowns
- **Credit limit -30%**: Reflect tightening
- **Utilization +25%**: Macro impact

6.2 Results

Scenario	Average Predicted EAD (kr)
Base	36,102
CCF +20%	36,997
Credit Limit -30%	29,957
Utilization +25%	36,218

Table 1: EAD under stress testing scenarios

7 Governance and Validation

7.1 Monitoring and Controls

- Regular R^2 and RMSE monitoring
- Quarterly backtesting and drift detection
- Challenger model periodically benchmarked

7.2 Model Risk and Limitations

- Model assumes behavioral stability over time
- Product-type encoding may proxy for segment effects
- Does not yet incorporate macroeconomic forecasts

8 Conclusion

- EAD model shows strong performance and logical behavior
- XGBoost captures nonlinear utilization and credit patterns
- Feature engineering and stress testing meet IRB expectations

Recommendation: Deploy model for internal EAD estimation and monitoring, with enhancements considered for future production use.