4th Seminar on Data Science and Analytics

Venue: Hotel Ramada Gurgaon Central

Date: 2nd November, 2019

Application of predictive models in the banking industry

Yash Ratanpal, AIAI

Manager – Acies Consulting LLP

Institute of Actuaries of India

Live Poll

Institute of Actuaries of India

Go to www.sli.do

Event Code: #7891

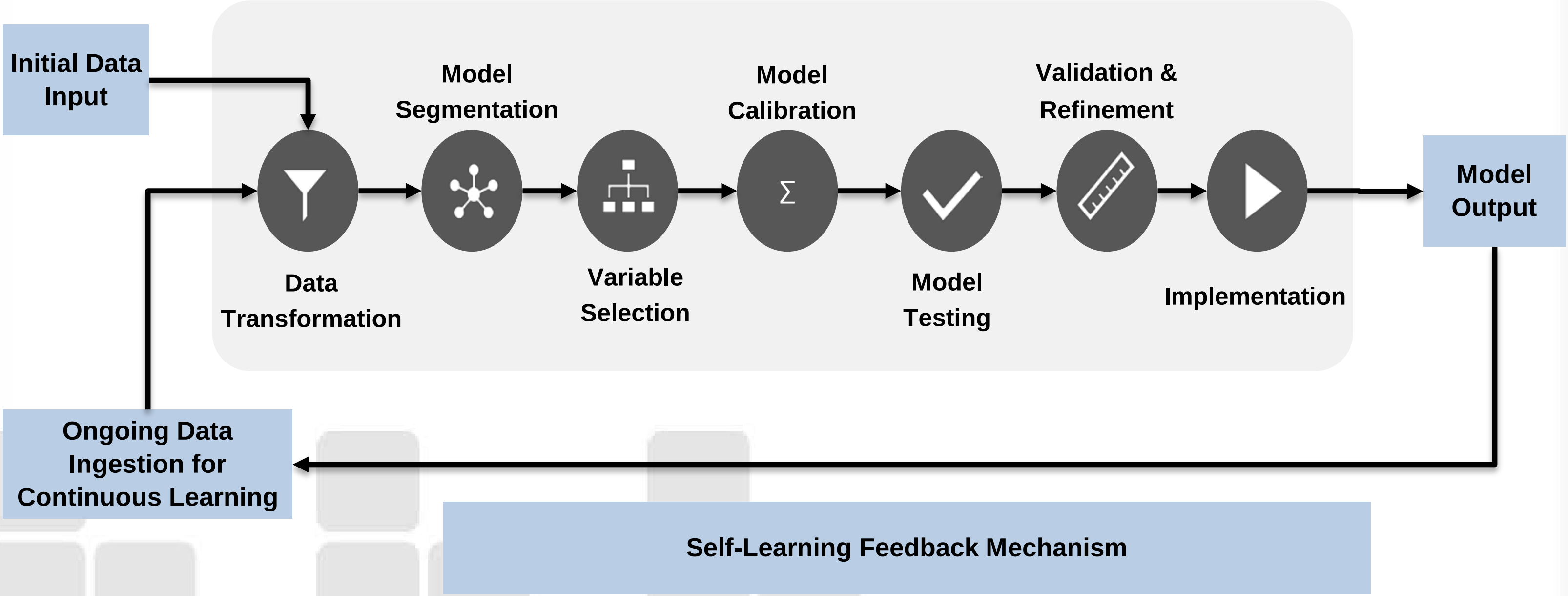
Table of Content

- A brief on predictive models
- Key drivers for increased usage of predictive models
- Application of predictive models in banking
- Commonly used predictive modelling techniques in banking
- Case study 1: Credit scoring models
- Case study 2 : Capital modelling and stress testing
- Key-takeaways
- Q/A

A brief on predictive models

"Predictive modeling is a statistical technique that works by analyzing historical and current data and generating a model to help predict future outcomes." – Gartner

The process of predictive modelling



Key drivers for increased usage of predictive models

1

Forward looking regulations

- Forward-looking regulations such as CCAR & DFAST, IFRS 9, IRRBB etc.
- Increased need for econometrics

2

Data growth

- Extensive centralization of structured data
- Exponential growth of unstructured data

3

Easier access to compute power

- Rise of open source technologies – Python and R
- Growth of cloud computing – Azure, AWS, GCP etc.

4

Increased desire for optimization

- Intensified scrutiny from stakeholders
- Desire for optimal performance and cost efficiency

Application of predictive models in banking

Institute of Actuaries of India

1 Credit Risk

- Loan underwriting & pricing
- ECL modelling (PD, LGD and EAD)
- Concentration risk
- Counterparty credit risk

2 Market Risk

- FX pricing
- Interest rate modelling
- Equity modelling
- Derivative pricing

3 Operational Risk

- Fraud Detection
- Legal loss modelling
- Catastrophic risk modelling

4 Liquidity Risk

- Liquidity stress testing
- Liquidity coverage ratio
- Net stable funding ratio

5 Asset-Liability Management

- Cashflow generation
- Behavioral modelling
- Fund transfer pricing

6 Other areas

- Integrated stress testing
- Regulatory and economic capital
- Cost optimization

Commonly used predictive modelling techniques in banking*

Institute of Actuaries of India

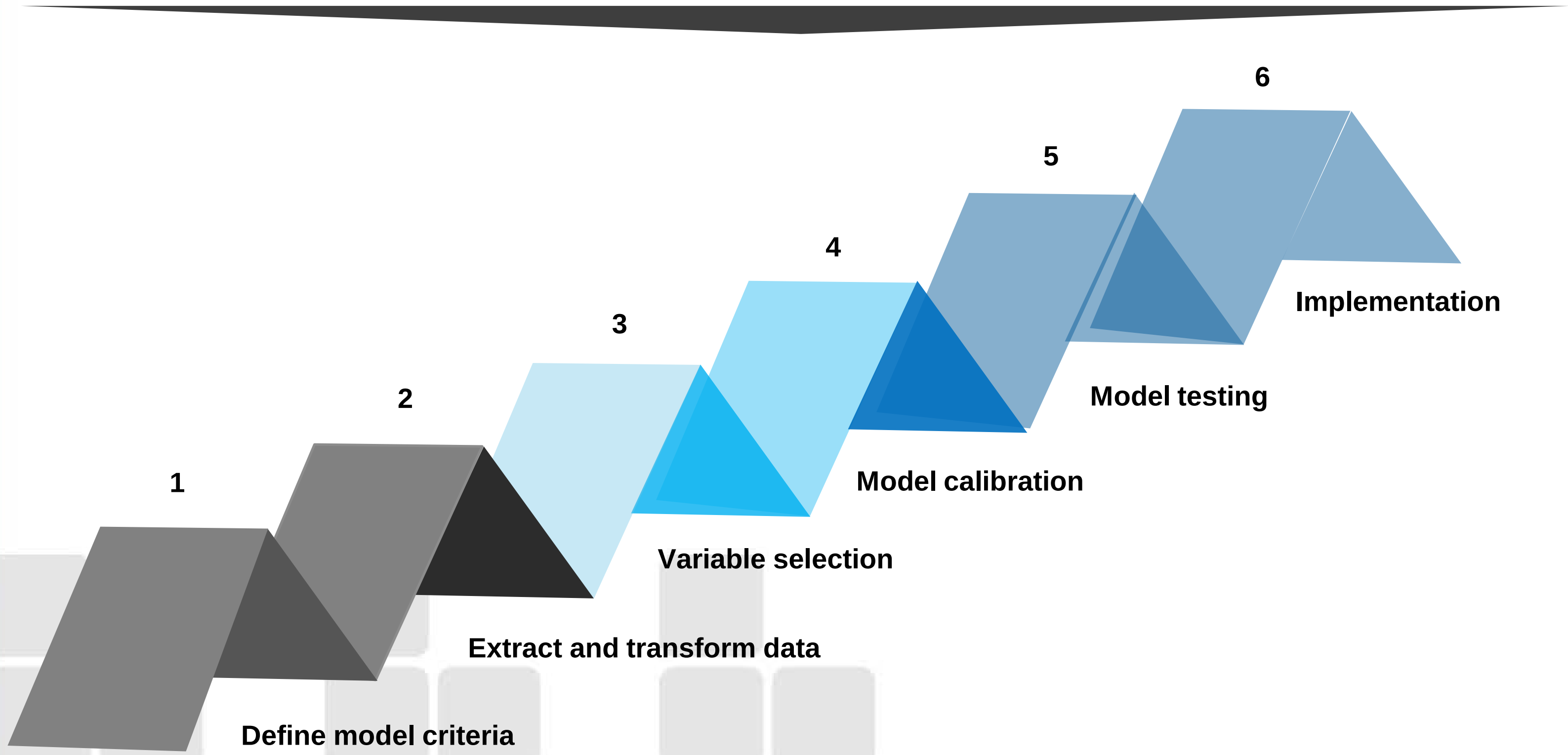
Technique	Credit Risk	Market Risk	Operational Risk	Liquidity Risk	Asset-Liability Management
Linear Regression		☑	☑	☑	☑
Logistic Regression	☑				
Time Series		☑	☑	☑	☑
Decision Trees	☑				
Random Forests	☑				
Neural Networks	☑				
PCA		☑			☑
K-Means clustering	☑	☑	☑	☑	
Gradient boosting	☑				

*Techniques and mapping shown above might not be comprehensive

Case study 1: Credit scoring models

A medium sized bank – XYZ, has recently developed and implemented a credit scoring model for its corporate credit portfolio. The credit scoring model is used to monitor credit health of existing corporate obligors and assess new corporate loan applications.

Process followed by bank to develop and implement the credit scoring model



Case study 1: Credit scoring models

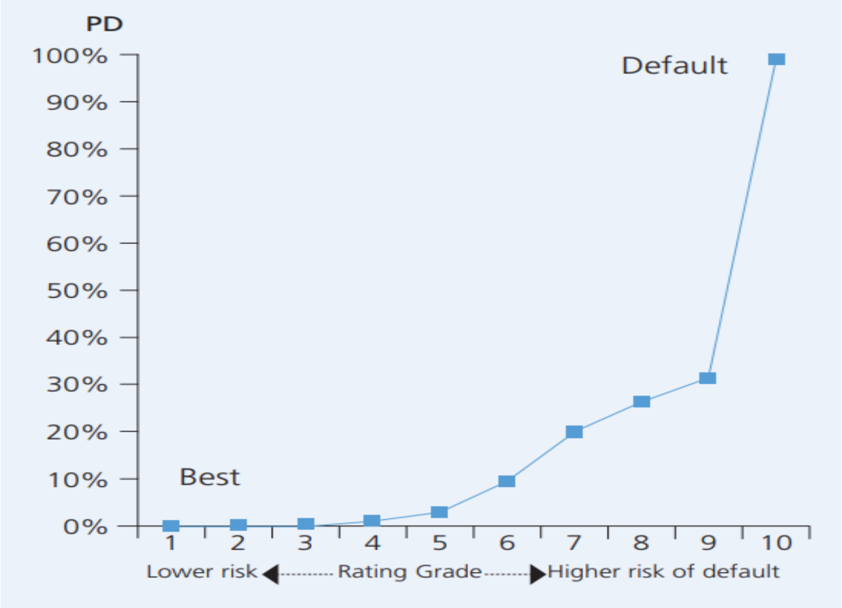
Institute of Actuaries of India

01

Define model criteria

- Model segmentation choice (Retail / Corporate / SME) – **Corporate**
- Required model outcome - **Number of rating outcomes to differentiate the risk of their borrowers. A range of 20 grades**
- Forecast horizon: **1-year PD**
- Default definition: **90 days past due (DPD)**

Simple illustration of model outcome



Case study 1: Credit scoring models

Institute of Actuaries of India

02

Extract and transform data

- Sample definition – Each rating grade requires a **sample of 1000 companies** and for **every 10 non-defaults, we need 1 default** in the sample
- Historical data period – Minimum period of **5 years**
- Type of data - financial measures such as p**rofitability, leverage, liquidity, cash flow adequacy** etc. obtained from **company financial statements**.
- Transformations – **missing values**, **outlier treatment, standardization** (log-value or ratios)

03

Variable selection

Quantitative variables

- Profitability
- Interest coverage
- Leverage
- Cash flows
- Liquidity
- Growth rates
- Operating cash
- flow/debt
- Debt/equity
- Profit margin
- Current ratios

Qualitative variables

- Reputation
- Track record
- Credibility
- Shareholding structure
- Industry profile
- Market position
- Management profile
- Shareholder support
- Parent strength

Case study 1: Credit scoring models

Institute of Actuaries of India

04

Model calibration

- Training vs testing dataset choice – **75 % of the sample data to be treated as training dataset while 25 % to be treated as testing dataset**
- Model framework – **Logistic regression framework**
- Measure used variable selection **– Information value criterion (IV)**

Logistic regression model framework

$$p = \frac{\exp(\beta_0 + \beta_1 \cdot x_1 + \cdots + \beta_n \cdot x_n)}{1 + \exp(\beta_0 + \beta_1 \cdot x_1 + \cdots + \beta_n \cdot x_n)}$$

where in the above

- p is the probability of default
- x_i is the explanatory factor i
- β_i is the regression coefficient of the explanatory factor i
- n is the number of explanatory variables

Case study 1: Credit scoring models

Institute of Actuaries of India

05

Model testing

- Measure used to assess discriminatory power of selected variables **– Information value criterion (IV)**
- Measures used to assess model accuracy **– GOF and Accuracy ratio**
- Measures used to assess discriminatory power of model **– ROC and confusion matrix**

Information Value (IV)

$$IV(x) = \sum_{i=1}^{N(x)} \left(\frac{g_i}{g} - \frac{b_i}{b}\right) \cdot log\left(\frac{\frac{g_i}{g}}{\frac{b_i}{b}}\right)$$

where,

- $N(x)$ is the number of levels in the variable x
- g_i represents the number of goods (no default) in category i of variable x_i
- b_i represents the number of bads (default) in category i of variable x_i
- g represents the number of goods (no default) in the entire dataset
- b represents the number of bads (default) in the entire dataset

Classification power	Information Value
Poor	<0.15
Moderate	Between 0.15 and 0.4
Strong	>0.4

Case study 1: Credit scoring models

Institute of Actuaries of India

Goodness of Fit (GOF)

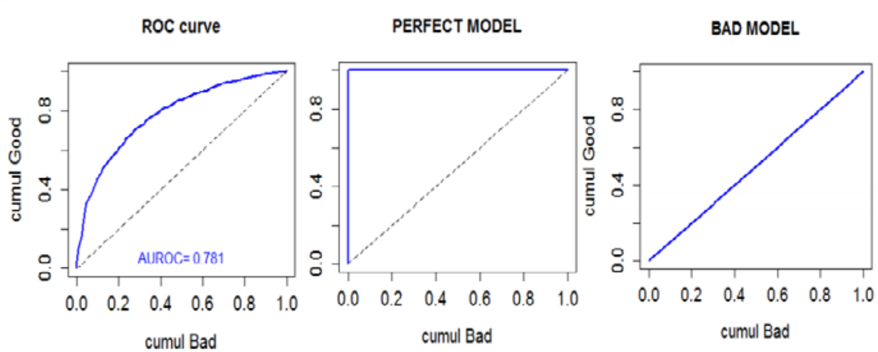
- The model predictions for every company in the test sample are grouped into say 10 percentile groups.
- Subsequently, the average PD is calculated for each group.
- Based on the total obligors in each group, derive the expected number of defaults per group.
- This is then compared to the actual number of realized defaults per group.
- This information is summarized for all 10 groups to derive the statistic that indicates how well the model predicts.

Company	Group	Predicted PD (by Model)	Average PD of Group	Expected no.of defaults (Group)	Realised no. of defaults (Group)	Accuracy Stat (Group)
A		1.00%				
B	1	2.00%	2.67%	0.08	0.01	85%
C		5.00%				
D		7.00%				
E	2	8.00%	9.00%	0.27	0.50	42%
F		12.00%				
:	3	:	:	:	:	:
:	4	:	:	:	:	:

Case study 1: Credit scoring models

ROC Curve

- ROC stands for "Receiver Operating Characteristic"
- The ROC curve is the plot between sensitivity and (1- specificity). (1- specificity) is also known as false positive rate and sensitivity is also known as True Positive rate.
- A perfect model would show a ROC curve that consists of two straight lines: From (0,0) to (0,1) and from (0,1) to (1,1), i.e. very steep.
- A model with no predictive power would have a ROC curve that follows the diagonal, since that would imply that for every cut-off value we find an equal number of goods and bads



Predictive Power	Area Under ROC
Acceptable	>70%
Good	>80%
Very Good	>85%

Case study 1: Credit scoring models

Institute of Actuaries of India

Confusion Matrix

- Used to test model's classification results against the actual observed classification.
- "True Positive Rate" corresponds to the fraction of Goods that are correctly classified [in the example below 7014/(7014+3171))
- "True Negative Rate corresponds to the fraction of Bads that are correctly classified (in the example below 357/(357+178)].

	Predicted Bad	Predicted Good
Observed Bad	357	178
Observed Good	3171	7014

Predictive Power	TP & TN rate
Acceptable	>60%
Good	>70%
Very Good	>85%

Case study 1: Credit scoring models

06

Implementation

- Instead of having distinct PDs for each borrower, XYZ devised **percentile groups** that **map to their internal rating scale** with a **number of 20 rating grades**
- Both **existing and new corporate borrowers** are evaluated with the model and given a rating (with corresponding PD) that is **commensurate with their credit profile**.

Mapping PDs to banks internal rating scale

Rating	PD Masterscale Risk Ranges Low	High
1. AAA	0.0000%	0.0119%
2. AA+	0.0119%	0.0168%
3. AA	0.0168%	0.0240%
4. AA-	0.0240%	0.0346%
5. A+	0.0346%	0.0500%
6. A	0.0500%	0.0733%
7. A-	0.0733%	0.1095%
8. BBB+	0.1095%	0.1636%
9. BBB	0.1636%	0.3033%
10. BBB-	0.3033%	0.5108%
11. BB+	0.5108%	0.8573%
12. BB	0.8573%	1.5114%
13. BB-	1.5114%	3.4872%
14. B+	3.4872%	5.9459%
15. B	5.9459%	7.7740%
16. B-	7.7740%	10.0928%
17. CCC+	10.0928%	12.5000%
18. CCC	12.5000%	15.0000%
19. CCC-	15.0000%	25.0000%
20. D	25.0000%	99.9999%

Each rating grade maps to a PD band. These PD can be re-calibrated when underlying risk profiles or lending conditions change.

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

As a response to the 2007-08 financial crisis, the federal reserve enforced CCAR – 'Comprehensive capital analysis and review', to provide an assessment of the health of systematically important banks in the United States and overall US banking sector.

Coverage
- CCAR risk charge covers retail banking, wholesale banking, trading, counterparty credit risk, operational risk, PPNR, fair value options and loans held for sale, AFS/ HTM securities.

Econometric Models
- Models used to derive forward-looking capital ratios need to be driven by economic intuition i.e. econometric models

Supervisory Stress Scenarios
- Forwarding looking capital ratios need to be computed under three supervisory scenarios – Baseline, Adverse and Severely Adverse

Thresholds for Breach
- BHCs are expected to maintain a minimum common equity tier-1 capital ratio of 4.5 under all supervisory scenarios

Outlier Banks
- Banks that breach minimum capital ratios would face ban on capital raising and distribution, severe fines & operational restrictions.

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

<u>Project Capital Ratios for each participating BHC – Severely adverse scenario</u>

Table 6. Projected minimum regulatory capital ratios in the severely adverse scenario, 2019:Q1 to 2021:Q1: 18 Participating Firms
Percent

Firm	Capital actions	Common equity tier 1 capital ratio		Tier 1 capital ratio		Total capital ratio		Tier 1 leverage ratio		Supplementary leverage ratio[1]	
		Actual 2018:Q4	Projected minimum	Actual 2018:Q4	Projected minimum	Actual 2018:Q4	Projected minimum	Actual 2018:Q4	Projected minimum	Actual 2018:Q4	Projected minimum
Bank of America Corporation	Original	11.6	5.6	13.2	7.1	15.4	9.4	8.4	4.5	6.8	3.7
	Adjusted										
The Bank of New York Mellon Corporation	Original	11.7	8.2	14.1	10.5	15.1	11.6	6.6	4.9	6.0	4.5
	Adjusted										
Barclays US LLC	Original	14.5	11.1	17.6	14.0	21.0	16.1	8.9	7.2	7.3	5.9
	Adjusted										
Capital One Financial Corporation	Original	11.2	3.9	12.7	5.5	15.1	7.7	10.7	4.8	9.0	4.0
	Adjusted	11.2	4.6	12.7	6.2	15.1	8.4	10.7	5.4	9.0	4.6
Citigroup Inc.	Original	11.9	6.9	13.5	8.4	16.6	11.2	8.3	5.2	6.4	4.0
	Adjusted										
Credit Suisse Holdings (USA), Inc.	Original	25.8	16.2	26.5	17.0	26.6	17.1	12.9	7.5	11.3	6.5
	Adjusted										
DB USA Corporation	Original	22.9	14.8	34.4	26.2	34.4	26.6	9.2	6.9	8.4	6.3
	Adjusted										
The Goldman Sachs Group, Inc.	Original	13.3	6.7	15.3	8.6	18.0	11.5	8.9	5.0	6.2	3.5
	Adjusted										
HSBC North America Holdings Inc.	Original	12.6	6.8	14.2	8.4	18.0	11.7	7.5	4.3	5.6	3.2
	Adjusted										
JPMorgan Chase & Co.	Original	12.0	4.4	13.7	6.3	15.5	8.3	8.1	3.8	6.4	3.0
	Adjusted	12.0	4.6	13.7	6.8	15.5	8.7	8.1	4.0	6.4	3.2
Morgan Stanley	Original	16.9	7.7	19.2	10.0	21.8	12.5	8.4	4.4	6.5	3.4
	Adjusted										
Northern Trust Corporation	Original	12.9	9.0	14.1	10.3	16.1	12.3	8.0	5.8	7.0	5.1
	Adjusted										
The PNC Financial Services Group, Inc.	Original	9.6	5.8	10.8	7.3	13.0	9.6	9.4	6.3	7.8	5.3
	Adjusted										
State Street Corporation	Original	11.7	8.2	15.5	11.8	16.3	12.5	7.2	5.5	6.3	4.8
	Adjusted										
TD Group US Holdings LLC	Original	16.3	12.4	16.3	12.4	17.3	13.6	9.2	7.1	8.3	6.4
	Adjusted										
UBS Americas Holding LLC	Original	21.7	11.0	25.7	16.6	27.0	18.6	11.3	7.2	n/a	n/a
	Adjusted									n/a	n/a
U.S. Bancorp	Original	9.1	6.5	10.7	8.2	12.6	10.4	9.0	7.0	7.2	5.6
	Adjusted										
Wells Fargo & Company	Original	11.7	7.0	13.5	8.6	16.6	11.7	9.1	5.8	7.7	4.9
	Adjusted										

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

A large US based BHC – XYZ, with consolidated assets more than $100 billion is preparing to submit their forward-looking capital action plan for the upcoming CCAR & DFAST cycle. The bank has used an inventory of more than 500 models to compute capital ratios under each of the supervisory scenarios.

Overview of XYZ's portfolio

Assets

- **Cash and cash equivalents**
- **Trading account assets**
- **Derivative assets**
- **Debt securities**
- **Loans and leases**
- **Fixed assets**
- **Receivables**
- **Goodwill**

Liabilities

- **Interest bearing deposits**
- **Noninterest bearing deposits**
- **Trading account liabilities**
- **Derivative liabilities**
- **Short term borrowing**
- **Long term debt**
- **Other liabilities**
- **Accrued expenses**

■ Line items with exposures more than $200,000 million

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

Major exposures in XYZs portfolio arising out of specific products were captured using specific models

Credit Risk	Counterparty risk	Sales and trading risk	Operational risk	Deposit Risk
Types of products: • Consumer real estate • Credit cards • Commercial real estate • Commercial lease financing	**Types of products:** • Interest rate derivatives • FX derivatives • Equity derivatives • Commodity derivatives • Credit derivatives • Debt securities	**Types of products:** • Interest rate products • FX products • Equity products • Credit products	**Types of products:** • All	**Types of products:** • Interest bearing deposits • Noninterest bearing deposits
Models used for CCAR: • Expected credit loss models (PD, LGD and EAD) • Rating transition models	**Models used for CCAR:** • CVA and DVA • VaR and expected-shortfall • Rating and credit spread models	**Models used for CCAR:** • Linear regression models - OLS • Time-series models – ARIMA	**Models used for CCAR:** • Modified loss distribution approaches • Copula based models	**Models used for CCAR:** • Money supply models • Co-integrated models for BSM & NII

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

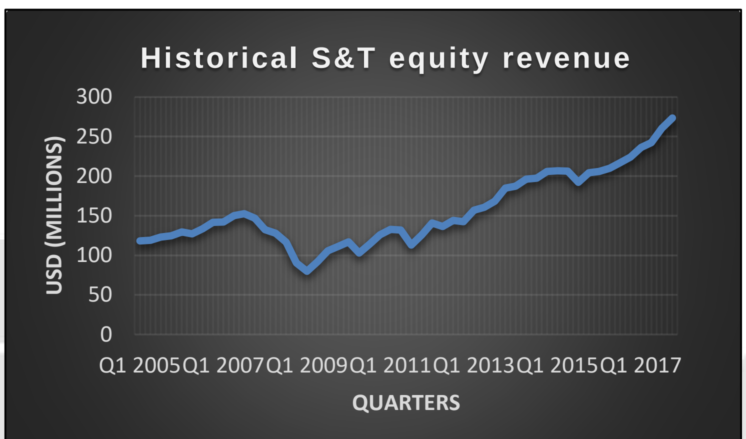
As part of this case study, we would be going through the end to end model development process followed by XYZ for a single PPNR model used to project sales and trading revenues pertaining to the equity asset class

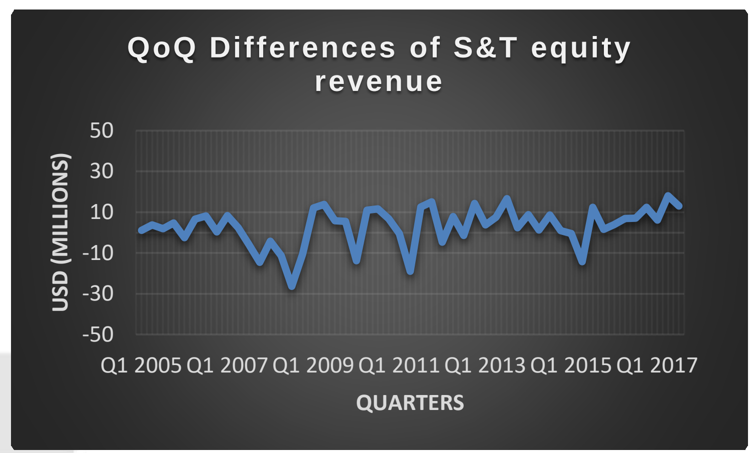
01 Defining the raw dependent variable

Historical time-series of the sales and trading revenue from the equity asset class

02 Create transformed dependent variable

Quarterly differences computed using the historical time-series of the raw dependent variable





Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

Use qualitative and quantitative analysis for variable selection

03 **Create repository of independent variables**

Productivity & Income

- U.S. Real GDP growth
- U.S. nominal GDP growth

Unemployment & Inflation

- U.S. unemployment rate
- U.S. CPI inflation:
- U.S. nominal disposable income growth
- Euro Inflation

Financial Markets

- U.S. 3-month Treasury rate
- U.S. BBB corporate yield
- Money Supply

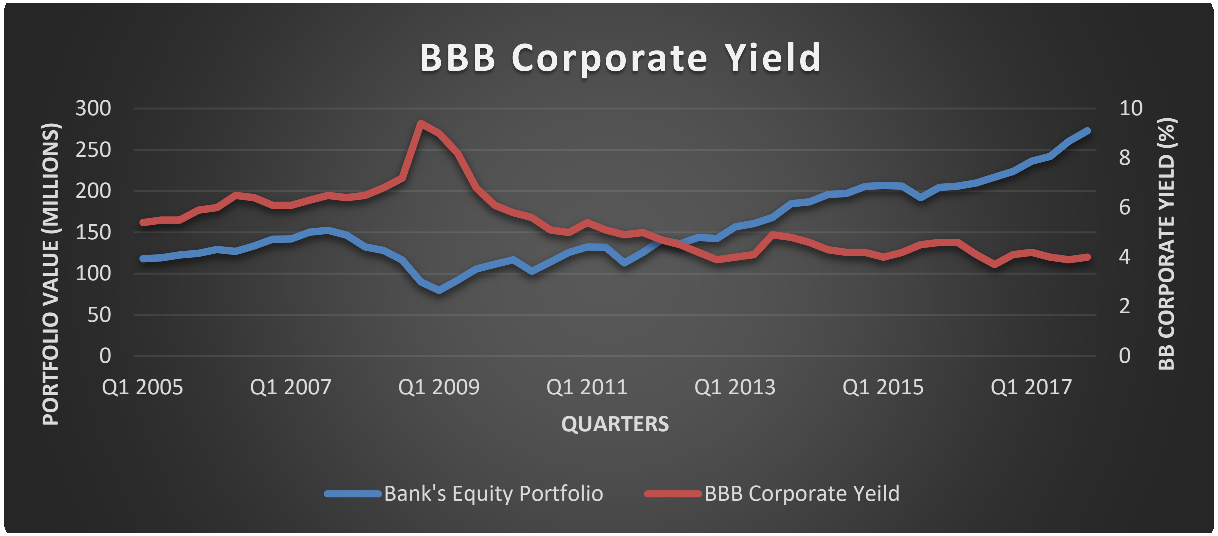
Indices

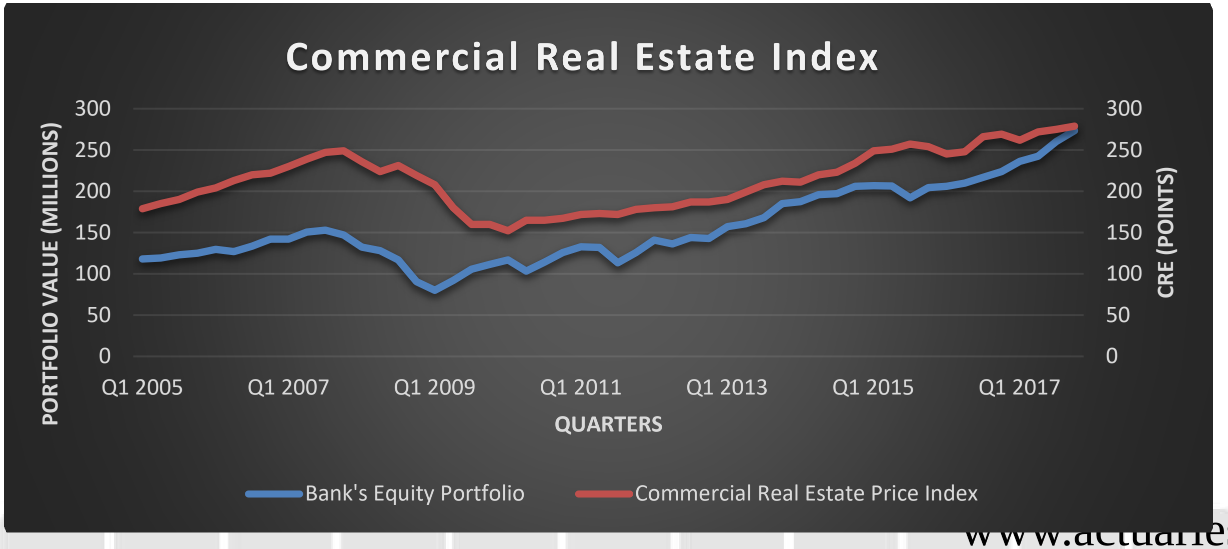
- S&P 500 and Dow Jones
- Commercial Real Estate Index
- Market Volatility Index (VIX)

Macro-Economic Variables	Intuitive Economic Relationship	Correlation
BBB Corporate Yield	Inverse	-72.50%
Commercial Real Estate Price Index	Direct	78.26%
Money Supply	Direct	85.82%
Consumer Confidence Index	Direct	80.83%
Business Confidence Index	Direct	63.48%
USD/EURO	Inverse	-61.54%

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India





Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

After identification of independent variables with the highest explanatory power and economic intuition, we proceed towards model estimation. OLS is a common estimation framework used for PPNR models.

04 **Model Calibration**

Model Equation

$$EquityPortfolio_t = \alpha + \beta_1 * ln(CRE_t) + \beta_2 * ln(BBBCorpYield_t) + \varepsilon_t$$

where:

α is the Intercept

β_1 is the regression coefficient of Commercial Real Estate Index

β_2 is the regression coefficient of BBB Corporate Yield

ε_i are residual errors (Actual – Predicted)

Economically intuitive model coefficients

Post-Calibration Results

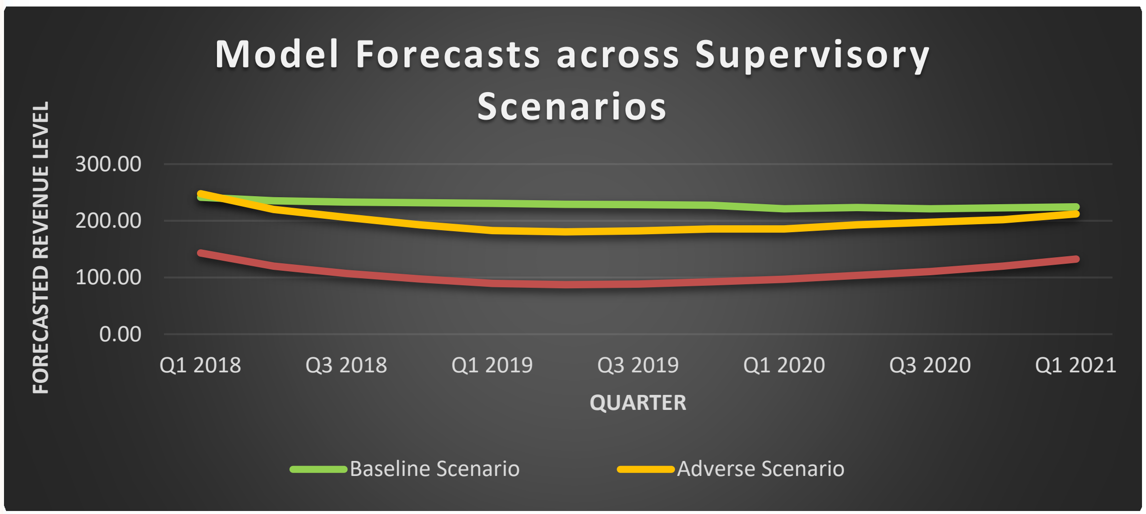
Parameter	*Coefficients*	*Standard Error*	*T-Stat*	*P-value*
Intercept	3.4451	0.3977	8.6625	< 0.0001
Log Transformed CRE	0.9744	0.0686	14.2132	< 0.0001
Log Transformed BBB Corporate Yield	-0.8172	0.0507	-16.1039	< 0.0001

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

Generate stress projections for CCAR & DFAST using projections of chosen independent variable as an input to the estimated model framework

5 **Generate stress projections for CCAR**



Model generates economically intuitive stress forecasts

Expected 9Q Loss Under Each Scenario (Millions)			
Initial Revenue Level (Q4 2017)	**Baseline Scenario**	**Adverse Scenario**	**Severely Adverse Scenario**
273.35	5.18	8.78	17.63

Case study 2: Capital modelling and stress testing

Institute of Actuaries of India

Stress projections generated by a diverse set of models covering different items of XYZs balance sheet – Credit losses, PPNR, operational risk losses, balances etc. are aggregated to compute required capital and leverage ratios for CCAR & DFAST compliance

6 **Computation of Capital Ratios**

Credit loss projections

PPNR projections

Deposit balances and NII projections

Counterparty and operational risk loss projections

Capital adequacy models, correlation models and RWA calculators

Projected regulatory capital and leverage ratios				
Regulatory Ratio	**Actual**	**Adverse Scenario (Minimum Ratios)**	**Severely Adverse Scenario (Minimum Ratios)**	**Minimum Ratio Threshold**
Common equity tier 1 ratio	10.4	7.5	4.4	4.5
Tier 1 capital ratio	12.8	8.7	6.9	6.0
Total capital ratio	13.3	10.9	9.1	8.0
Tier 1 leverage ratio	8.1	5.5	3.9	4.0
Supplementary leverage ratio	6.4	4.6	3.5	3.0

Key-Takeways

Institute of Actuaries of India

Forward-Looking Regulations

- Ability to meet new forward-looking regulations
- CCAR, IFRS 9, FRTB, IRRBB etc.

Access to data and compute power

- Exponential growth of structured and unstructured data
- Rise of open source technologies and cloud compute

Desire for optimization

- Intensified scrutiny from stakeholders
- Need for cost, performance and capital optimization

Actuaries & Predictive modelling

- Actuaries skillset compatible with predictive modelling
- Combination of statistics, business, finance and economics

Q/A

Institute of Actuaries of India

Questions ?

Comments

Thank You

Application of predictive models in the banking industry

Yash Ratanpal, AIAI

Manager – Acies Consulting LLP

Institute of Actuaries of India