SAILES

Schneider Electric Challenge

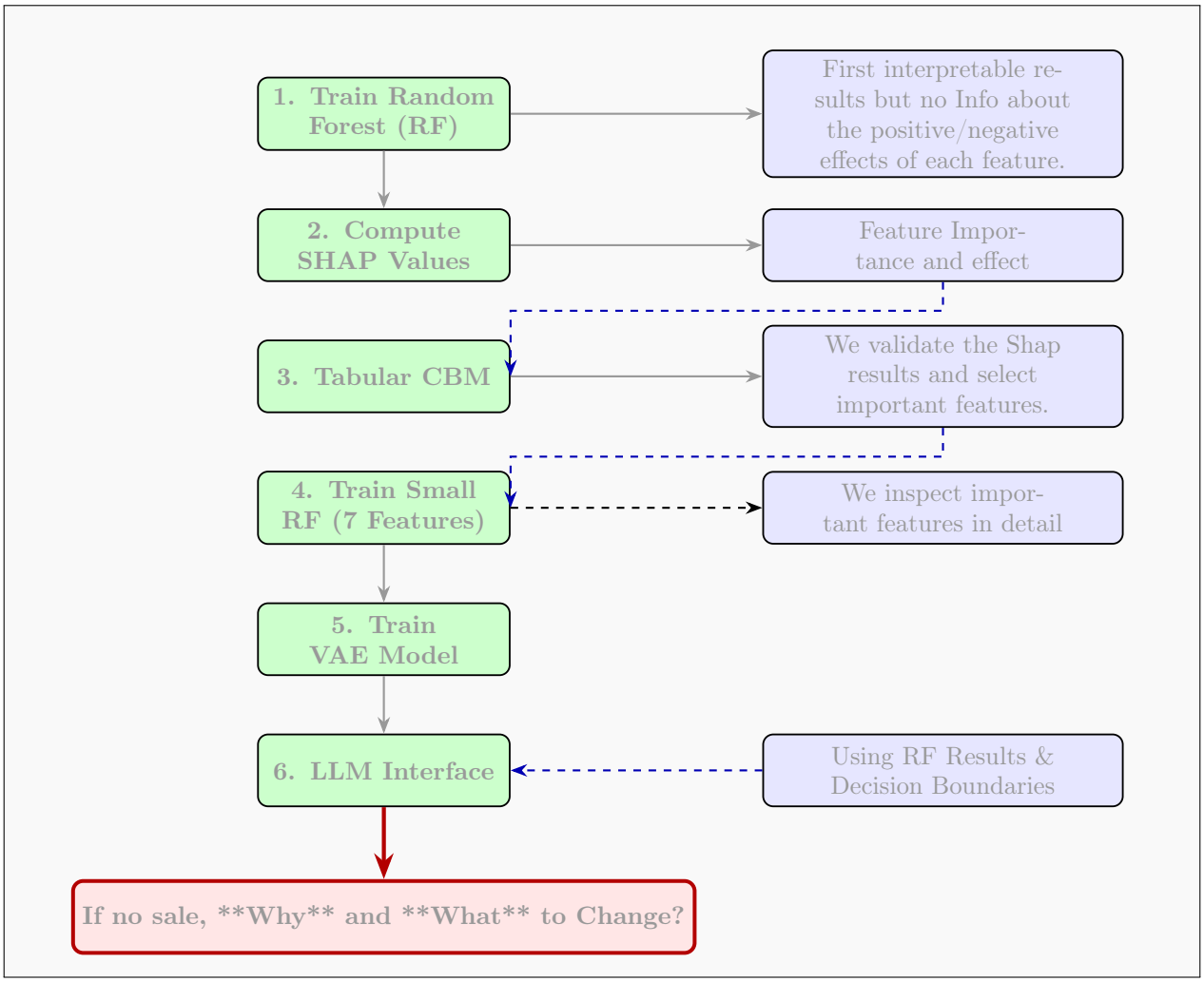


Figure 1: Our project Methodology Flowchart

Authors: Anooj Sathyan | Jaycent Gunawan | Marcello Galisai | Guillem Masdemont

Contents

1 Who we are!

Hello! We are an international team of four, Anooj Sathyan from India, Jaceny Gunawan from Indonesia, Marcello Galisai from Italy and Guillem Masdemont from Spain.

1.1 The Challenge

In sales, understanding *why* we win or lose opportunities and *what to change* to obtain future wins is as important as predicting the outcome itself. This project tackles a critical business question:

"Why does our AI model predict that a sales opportunity will be won or lost?"

1.2 What We Did

We developed and analyzed multiple machine learning models to predict sales outcomes, then applied cutting-edge explainability techniques to make these predictions **transparent** for sales teams, managers, and executives.

Key Insight

We can now explain *every single prediction* in simple and task-grounded terminology. For example: "This opportunity has a high probability of success because the customer has a 95% historical success rate and we've had 12 previous positive interactions with them."

We developed SAILES, an LLM-powered assistant tool, to leverage the predition models and provide natural language insights to the sales team.

1.3 Business Impact

- **Sales Teams:** we allow consulant to focus efforts on high-impact actions by understanding which factors matter most
- **Strategy:** we enable to identify systemic patterns in wins/losses to refine go-to-market approach
- **Customer Success:** we proactively uncover risk factors before opportunities are lost

2 The Business Problem

2.1 Context

Schneider Electric generates millions of data points from sales opportunities stored in our CRM system. Each opportunity represents a potential sale, with information about:

- **Customer History:** Past purchases, success rates, number of contracts
- **Product Mix:** Which products we're trying to sell and in what quantities
- **Competitive Landscape:** Which competitors are involved
- **Timing & Relationship:** How long the opportunity has been open, interaction frequency

2.2 The Challenge

Traditional machine learning models act as "black boxes" – they make predictions but don't explain their reasoning. For business users, this creates several problems:

1. **Trust:** Hard to trust predictions you don't understand
2. **Action:** Can't take corrective action if you don't know what factors matter
3. **Learning:** Can't improve processes without understanding success patterns
4. **Accountability:** Difficult to explain decisions to customers or management

Actionable Recommendation

Developing an LLM-powered solution, buildingo on highly accurate and fully-interpretable predictions.

3 Our Approach: Three-Layered Explainability

We implemented a comprehensive explainability framework using three complementary techniques:

Global Insights	Local Explanations	What-If Analysis
What matters overall?	Why this specific prediction?	What would change it?

Figure 1: Three Pillars of the Interpretability Framework

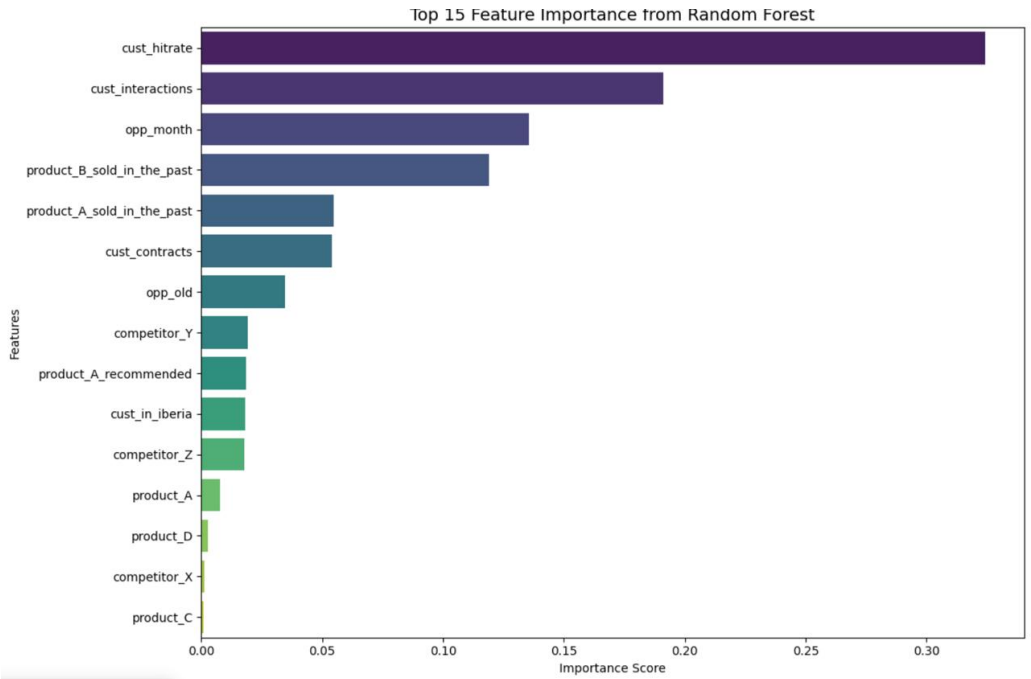
3.1 Layer 1: Global Insights – The Big Picture

What it tells us: Which factors are most important across all sales opportunities?
Technique Used: Feature Importance Analysis and Concept-Based Models (TabCBM)

Technical Details

Feature Importance: Top 5 Most Important Factors for Winning Opportunities

1. **Customer Hit Rate**: Historical success rate with the customer
2. **Number of Interactions and Contracts**: How much engagement and business relationships we've had
3. **Historical Sales**: Whether we've successfully sold a product before
4. **Number of Interactions**: How much engagement we've had
5. **Novelty of the opportunity**: Is the opportunity relatively long-standing? → Long-standing offers tend to be less successful



Technical Details

TabCBM: Concept Analysis
TabCBM is a cutting-edge AI system, that automatically extracts most relevant concepts (i.e., groups of feature that incapsulate a lot of information) from tabular datas. The results are clear: the decisive factor is the costumer and the relationship we have with them. In particular, costumer with many past interactions, many past contracts and relatively novel opportunities strongly influence the wins.

What this means in practice:

- Strong existing relationships are the biggest predictor of success
- Past success breeds future success
- Regular customer engagement significantly improves win rates
- Competitor presence has less impact than expected

This is information constitutes part of background knowledge of the LLM-assistant, allowing it to interpret information based on our global findings.

3.2 Layer 2: Local Explanations – Individual Predictions

Technical Details

Technique: Decision Path Analysis
What it tells us: Why did the model make this specific prediction for this particular opportunity?
The LLM-assistant is powered with the results from Decision Path Analysis, i.e. data-driven analysis of which specific information had the highest impact to produce the prediction at hand. How is this possible? The prediction model (a Random Forest) is composed of several parallel sub-models, specialising in the analysis of certain trends in the data: Decision Path Analysis recovers information about which of these models acted as "tie-breakers" and, therefore, which trends where more important.

3.3 Layer 3: What-If Analysis – Counterfactual Explanations

Technical Details

Technique: Tabular Variational AutoEncoder (VAE)
What it tells us: What would need to change for this prediction to flip?
Tabular Variotional AutoEncoder are state-of-the-art deep encoders, allowing to identify hidden and deep information and uncovering similarities among data points. This can be exploited in the following way: given an opportunity negative prediction, the VAE produces the *most similar plausible opportunity* that would have produced a positive prediction, providing the minimal amount of feature changes possible.

Actionable Recommendation

Why this matters: Actionable Insights
Sales teams can now see exactly what actions may have a big impact to overturn a possible loss, allowing them to prioritize efforts on the highest-leverage activities.

4 Model Performance

4.1 Technical Metrics

Model	Accuracy	F1 Score	Precision	Recall
Random Forest	0.83%	0.81	0.83	0.83

Model	Validity Rate	Confidence	Changed Features	Proximity
Tabular VAE	100%	0.73	1.5	1.2

4.2 Interpreting results

- **83% Accuracy:** Out of 100 opportunities, we correctly predict 83
- **83% Precision:** When we predict a win, we're right 83% of the time
- **83% Recall:** We identify 83% of opportunities that will actually be won
- **100% Validity, 0.73 Confidence, 1.2 Proximity:** The VAE produces highly plausible counterfactuals
- **1.5 Changed Features:** The VAE provides precise and direct actions to flip predictions

5 Key Findings & Actionable Insights

5.1 Finding 1: Relationship Quality Dominates

Key Insight

Customer relationship metrics (hit rate, interactions, contracts) account for the **majority** of the prediction power, while competitive factors are **minimal**.

Actionable Insights:

- **For Sales Teams:** Prioritize relationship building and frequent touchpoints over purely product-focused pitches
- **For Management:** Invest in CRM tools and processes that track and encourage regular customer engagement
- **For Strategy:** Consider relationship quality as a key criterion in opportunity qualification

5.2 Finding 2: Past Product Success is a Strong Predictor

Key Insight

Customers who have previously purchased Product A are **more likely** to win new opportunities compared to new-to-Product-A customers.

Actionable Insights:

- **For Sales Teams:** Prioritize cross-selling and upselling to existing Product A customers
- **For Product:** Focus on customer success and retention for Product A to drive future sales
- **For New Customer Acquisition:** Develop specific strategies (pilots, demos, case studies) to overcome the "first sale" hurdle

5.3 Finding 3: Stalled Opportunities Rarely Recover

Key Insight

Longstanding opportunities have a |**lower** win rate, and the impact compounds over time.

Actionable Insights:

- **For Sales Teams:** Implement 60-day urgency checkpoints; if no progress, either escalate or close
- **For Management:** Track "opportunity age" as a key pipeline health metric
- **For Process:** Create structured "deal rescue" playbooks for opportunities at risk of stalling

6 Technical Innovation

6.1 Why Our Approach is Advanced

While maintaining the focus on business usability, our solution incorporates several technical innovations:

1. **Concept-Based Modeling:** Unlike traditional neural networks, TabCBM learns human-interpretable concepts that naturally map to business terminology
2. **Multi-Method Explainability:** Combining feature importance analysis, decision path analysis, concept analysis, and counterfactuals provides comprehensive understanding
3. **Counterfactual Generation:** VAE-based approach generates realistic "what-if" scenarios that are actually achievable
4. **Sparsity Optimization:** Focused on minimal changes (average 3-4 factors) to make recommendations actionable

6.2 Comparison to Standard Approaches

Aspect	Traditional ML	Our Approach
Prediction Only	Yes	Yes + Why + What-if
Explanation	None or post-hoc	Built-in from model architecture
Actionability	Limited	Specific recommendations
User-Friendliness	Technical metrics only	Plain language insights
Trust	"Black box"	Transparent reasoning

Table 1: Our Approach vs. Traditional Machine Learning

7 Limitations and Future Work

7.1 Current Limitations

We're transparent about what the model can and cannot do:

- **Historical Bias:** Model learns from past data, so it reflects historical patterns (good and bad)
- **Causation vs. Correlation:** Model identifies patterns but doesn't prove causation
- **Data Quality Dependent:** Predictions are only as good as the CRM data quality
- **Static Factors:** Some factors (customer industry, size) can't be changed

7.2 Future Enhancements

- **Real-Time Updates:** Recalculate predictions as opportunities evolve
- **Automated Alerts:** Proactive notifications when opportunities become at-risk
- **Integration:** Connect with email and calendar for automatic interaction tracking

8 SAILES: Supervised Artificial Intelligence Learning Explainable System

8.1 Presenting our interface

This interface presents the **SAILES** (Supervised Artificial Intelligence Learning Explainable System) dashboard, designed to provide transparent, explainable insights into our model's prediction of whether an opportunity will be Won or Lost. On the left, users input **Customer Metrics** and **Opportunity Details** to generate a prediction. The central panel provides high-level model explanation by answering crucial business questions (e.g., "Which competitors matter most?"), detailing how specific factors influence the outcome in a moderate-negative direction. This setup allows business users to understand the model's general strategy and apply the derived insights to real-world sales scenarios.

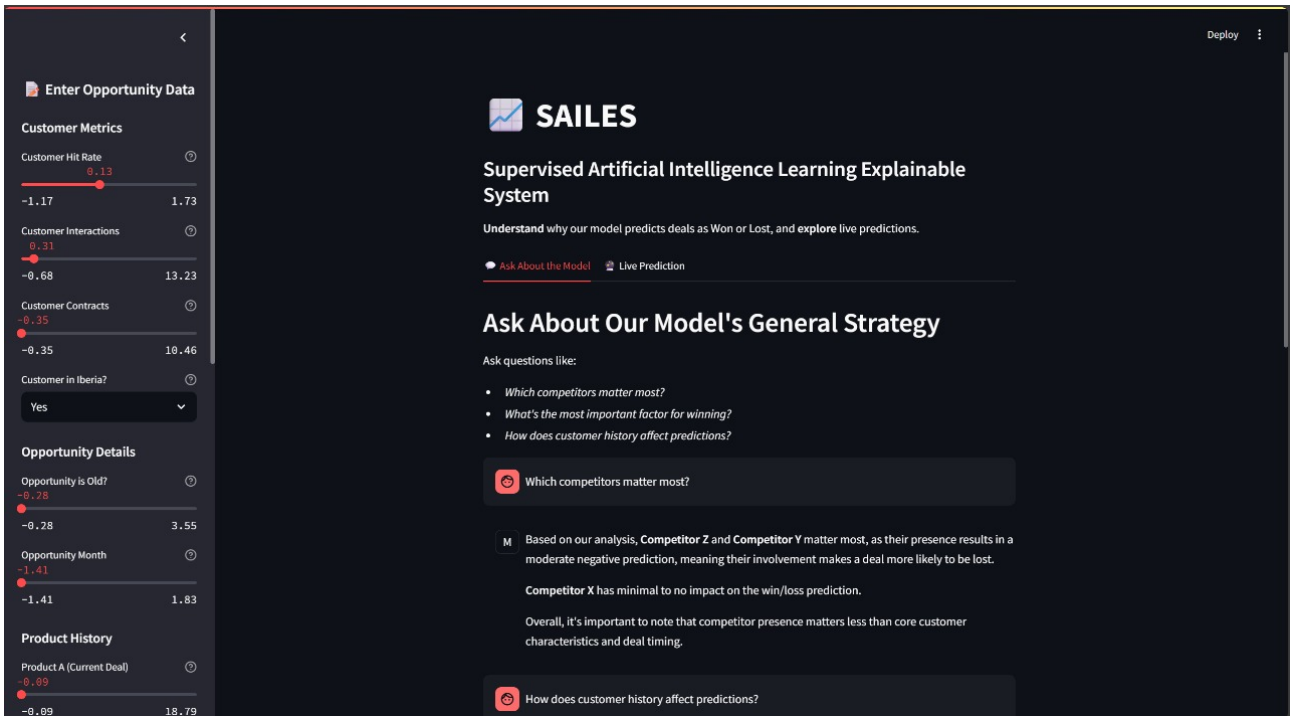


Figure 2: The SAILES (Supervised Artificial Intelligence Learning Explainable System) dashboard, illustrating how Customer Metrics and Opportunity Details are used to generate predictions, with the central panel providing plain-language explanations to critical business questions.

A Resources

1. GITHUB REPOSITORY:

   Link: GitHub repository.

2. TECHNICAL REPORT:

   Link: Technical report.

3. SAILES MVP
   The Link to the MVP is in the GITHUB repository!