

Behavioral Early Warning System for Credit Defaults



Polto Fabio

Head Monitoring & Credit Quality @Banca Sella

Stefano Priola

CTO @Centrico

Federico Angaramo

Data Scientist @Centrico

CIPA Workshop - January 29, 2021

Banca
Sella

{ centric :
co

Agenda

- Old Early Warning Model : Introduction to CADR
- Old Early Warning Model: Critical points
- New Behavioral Early Warning System
- Results & Lesson Learned
- Next steps

Old Early Warning Model

Introduction to CADR

Contesto di riferimento

- In Banca Sella esistono diverse procedure di rilevazione del rischio di credito.
- Una di esse è la procedura di Early Warning Cadr
- Si tratta di una procedura nata nel 2007 con lo scopo di individuare il rischio di credito di brevissimo periodo (0-3 mesi) basandosi essenzialmente sull'esperienza dei gestori dei crediti anomali nell'interpretare i fattori che abitualmente sono premonitori di un default.

Processo di monitoraggio del rischio di credito: la vecchia procedura di Early Warning Cadr

- Scopo della procedura era assegnare ad ogni posizione affidata o in utilizzo un colore Cadr che fosse rappresentativo del rischio di credito di brevissimo periodo.
- In ordine di gravità i colori erano: verde, giallo, rosso, nero.
- In Banca Sella sono sottoposte e assegnate a un gestore specializzato tutte le posizioni in bonis che presentano pessimi indicatori di rischio.
- Ai fini del presidio, tutte le posizioni in Cadr nera sono oggetto di monitoraggio mentre le posizioni in Cadr rossa o gialla possono essere sottoposte a monitoraggio in base all'entità del credito
- La tempestiva attribuzione del corretto colore Cadr è fondamentale per una efficace azione preventiva finalizzata alla riduzione del rischio di credito.

Old Early Warning Model

Critical points

Criticità del vecchio modello Early Warning CADR in uso dal 2007 presso Banca Sella:

- ✓ La vecchia procedura Cadr era suddivisa in 5 principali Gruppi di indicatori: sconfini, rate, anticipi, rischi esterni, rating . Ad ogni Gruppo di indicatori veniva assegnato un colore, da verde a nero; il colore complessivo era basato su un sistema di punteggi. Ad ogni anomalia rilevata veniva attribuito un punteggio stabilito su base esperienziale.
- ✓ La vecchia procedura Cadr utilizzava un modello unico valido per tutti i segmenti di clientela senza differenziare aziende e privati, crediti rateali o non rateali, utilizzi grandi e utilizzi piccoli
- ✓ La suddivisione della clientela in base al colore Cadr non seguiva una distribuzione stabile e funzionale al monitoraggio
- ✓ La Cadr non sempre era sufficientemente predittiva: in alcuni casi il colore Cadr diventava nero dopo l'insorgere di anomalie ormai gravi

Requisiti per nuovo Early Warning CADR

La banca aveva bisogno di un nuovo modello di Early Warning che soddisfacesse 4 requisiti:

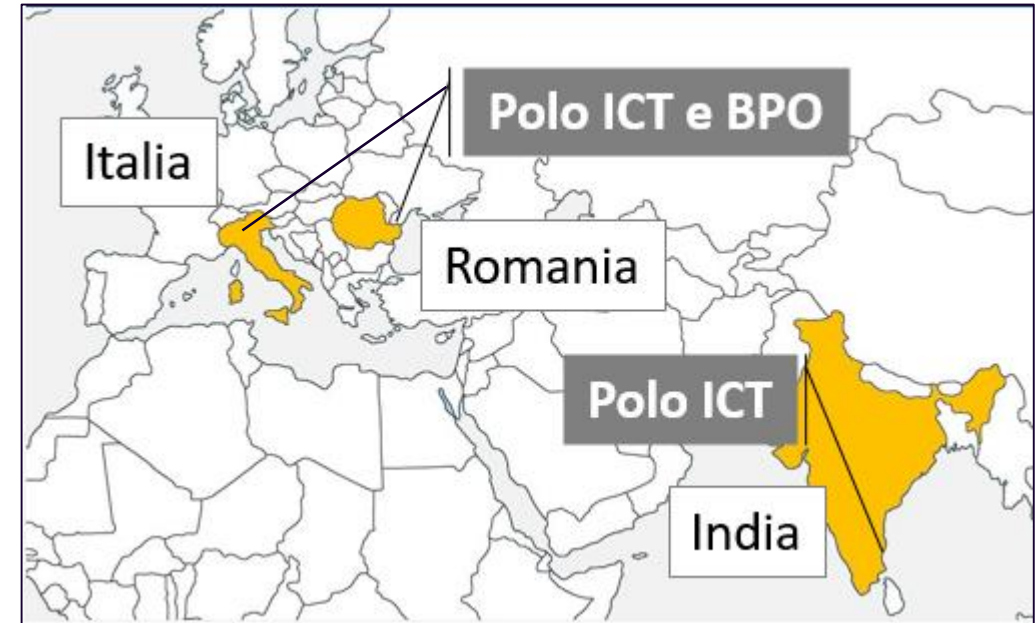
- 1) **Maggior tempestività:** i crediti a rischio default devono essere intercettati e analizzati con un preavviso superiore ai 3 mesi rispetto al data di presunto deterioramento. Ciò al fine di poter porre in essere eventuali misure per ridurre il rischio
- 2) **Output più funzionale:** la distribuzione della classificazione da modello Cadr deve necessariamente essere funzionale alla reale rischiosità del portafoglio e alle capacità di monitoraggio delle strutture di presidio.
- 3) **Probabilità di default più accurata,** da ottenersi con un sistema di punteggi, o un modello, che tenesse conto della reale rischiosità delle variabili misurate sul debitore
- 4) **Coerenza con il rating:** la probabilità di default di breve periodo, espressa dalla procedura di early warning, non poteva essere migliorativa della Pd 12 mesi o Lifetime, soprattutto in previsione di passaggio al nuovo sistema di rating Airb

New Behavioral Early Warning System

Who we are

Who is Centrico?

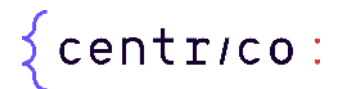
- Operational and IT backbone of Sella Group since 1968
- Starting from 2019, Open Banking outsourcer focused on bank and fintech core banking
- IT and BPO full outsourcing supported by 600 HC in Italy, 400 HC in Romania and 300 HC in India (900 IT and 400 BPO)



1,300 people in 3 countries

Some numbers and facts:

- Open technology (mainly java based) modular core banking without legacy dependencies
- >1,700 server in private cloud
- ~ 3 millions of accounting transactions per day
- > 1 billion API calls yearly
- ~ 200 (attended and unattended) RPA & AI flows



We enable banks and
fintechs to innovate
thanks to our proven
open banking system,
that is easy to integrate
and customise,
and minimises time to
market



New Behavioral Early Warning System

Solution overview

High Level Solution

Design

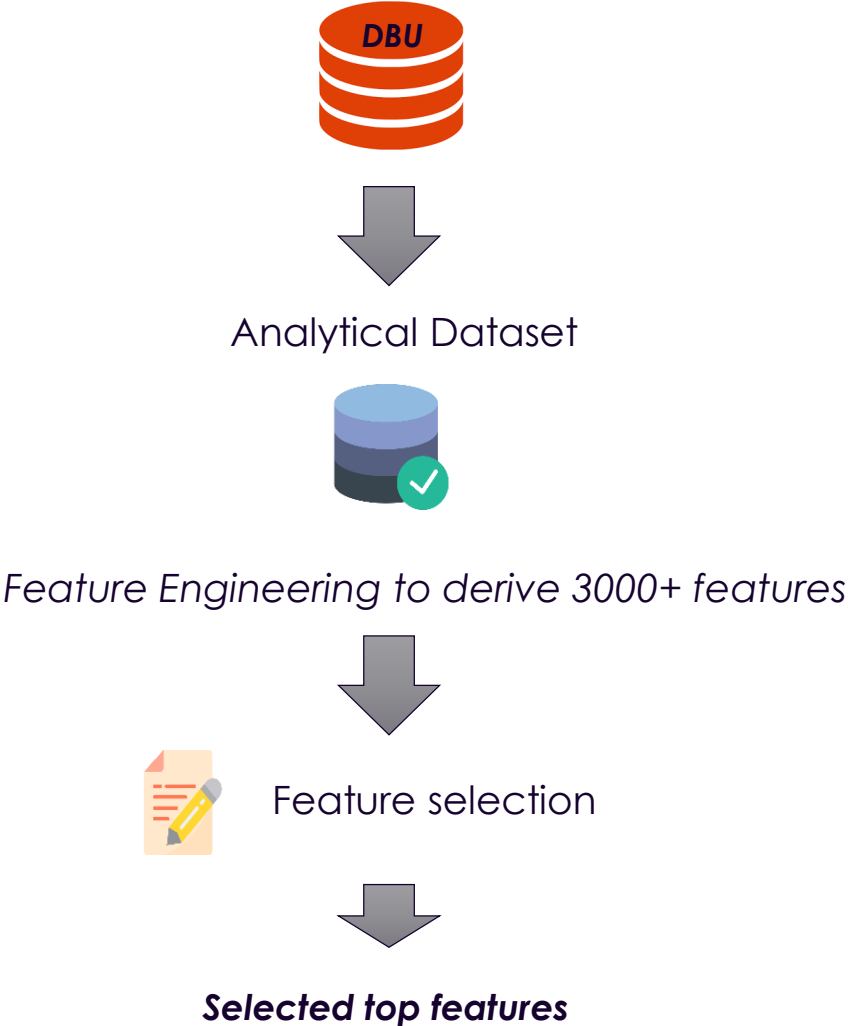
Products in Scope:

 Credit lines

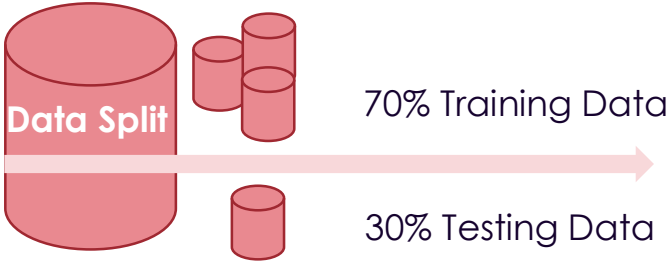
Customers in Scope:

-  SME / Corporate
-  Ditte individuali
-  Privati

Data preparation



Modeling



Algorithm Choice



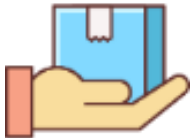
Model Evaluation

Model accuracy measurement through defined KPIs like

- Hit Rate
- AUC

Feature Relevance
OOT validation

Deployment



Snapshot creation automation



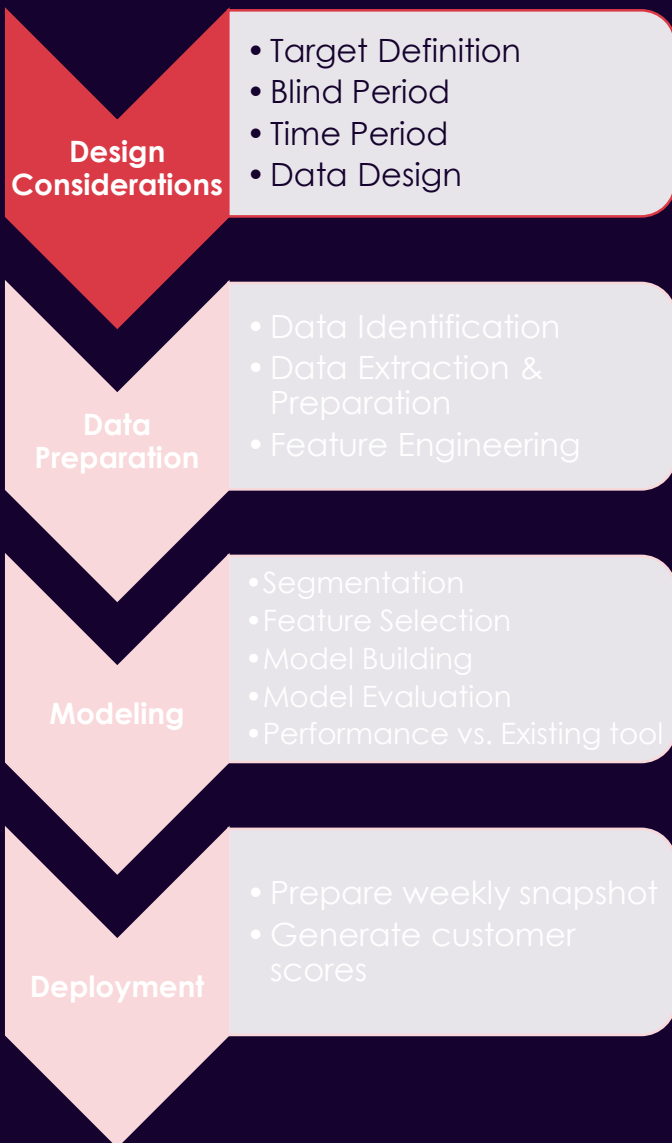
Model running automation



Periodic back-testing for model relevance and validation

New Behavioral Early Warning System

Solution details



Design Considerations

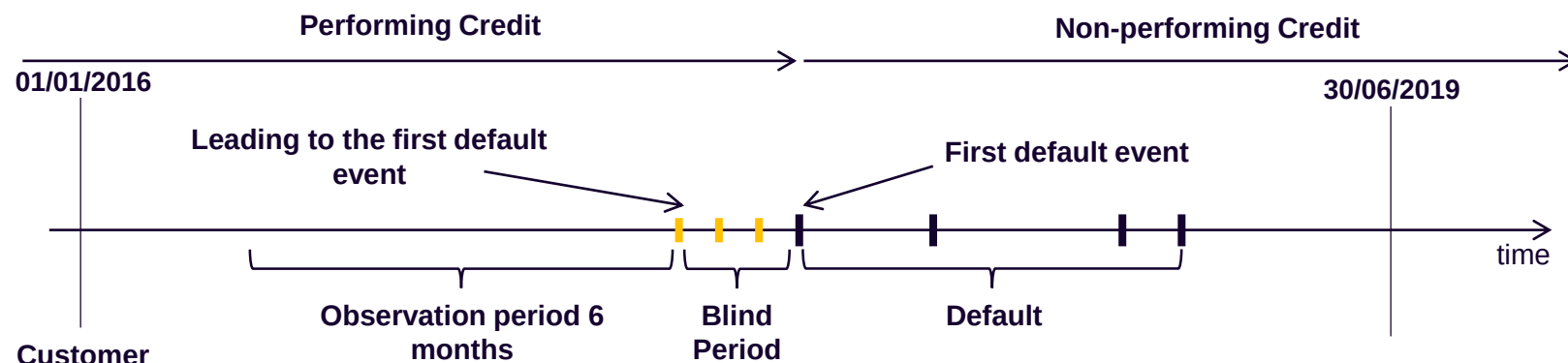


Target Definition

- ✓ Main default events and simulated events
- ✓ Included Strong and weak default events (payment past due, default due to forbearance etc.)
- ✓ If client had the very first default event 3 months from given point in time (Yes/No)



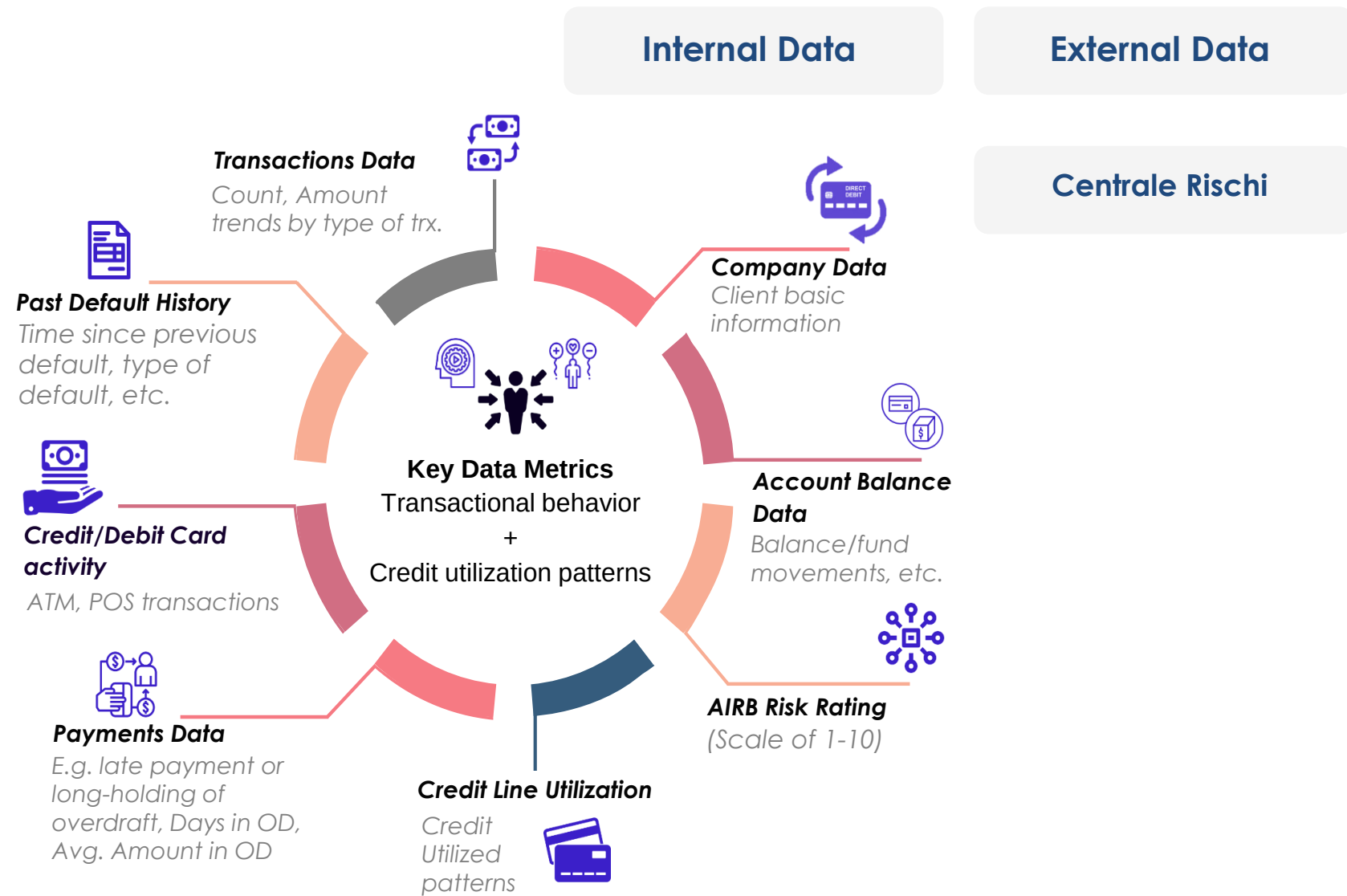
- ✓ **Blind period** (number of months in advance that an event should be predicted) -> 3 months
- ✓ **Observation period** -> 6 months



Note: the observation is always related to a performing period



Data Identification





Data Extraction & Preparation

- ✓ Data Extracted at daily/weekly/ monthly/ quarterly levels

- ✓ Aggregate to get weekly level data
- ✓ Calculate trends, ratios etc.
- ✓ Create snapshots based on previous 26 weeks



Raw Data



Aggregation



Data Preparation



Snapshot Generation

Filters and Business Exclusions applied:

- Only clients already under the existing scoring system were considered



Feature Engineering



Base Features:

- Examples Exposure Amount, Transaction Amount, Account balance amount, etc

Directly obtained from raw inputs without any modification.



Velocity Variables:

- Example: Count of transactions less than €50 in last 4 weeks / last 26 weeks, last 4 weeks / first 12 weeks, sum of last 8 weeks etc.

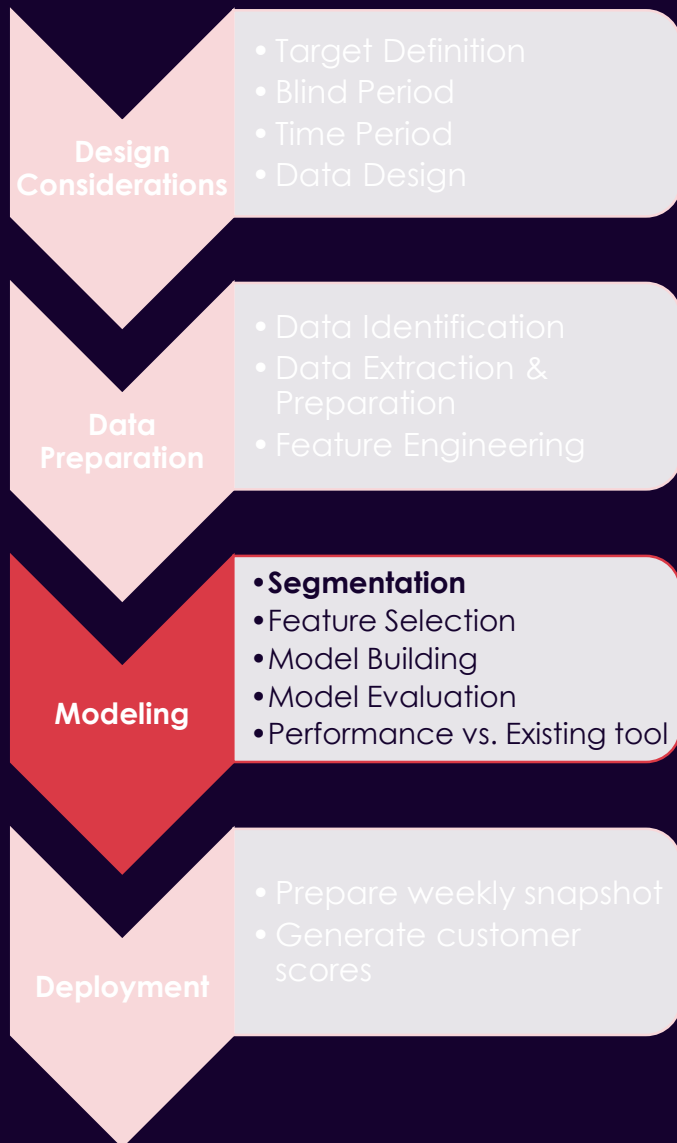
Trends/ratios derived from Base Features.



Discrete Variables:

- Example: Customer accorded, Utilized trends, credit lines

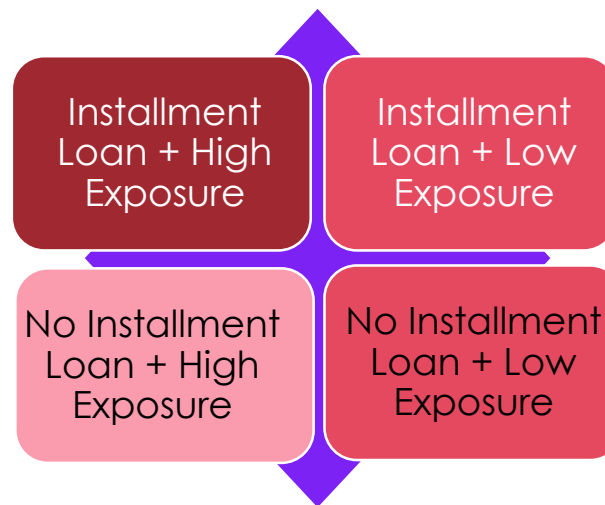
Created by binning or grouping the raw variables.



Segmentation

For each type of customer

Based on type of credit* and exposure level*

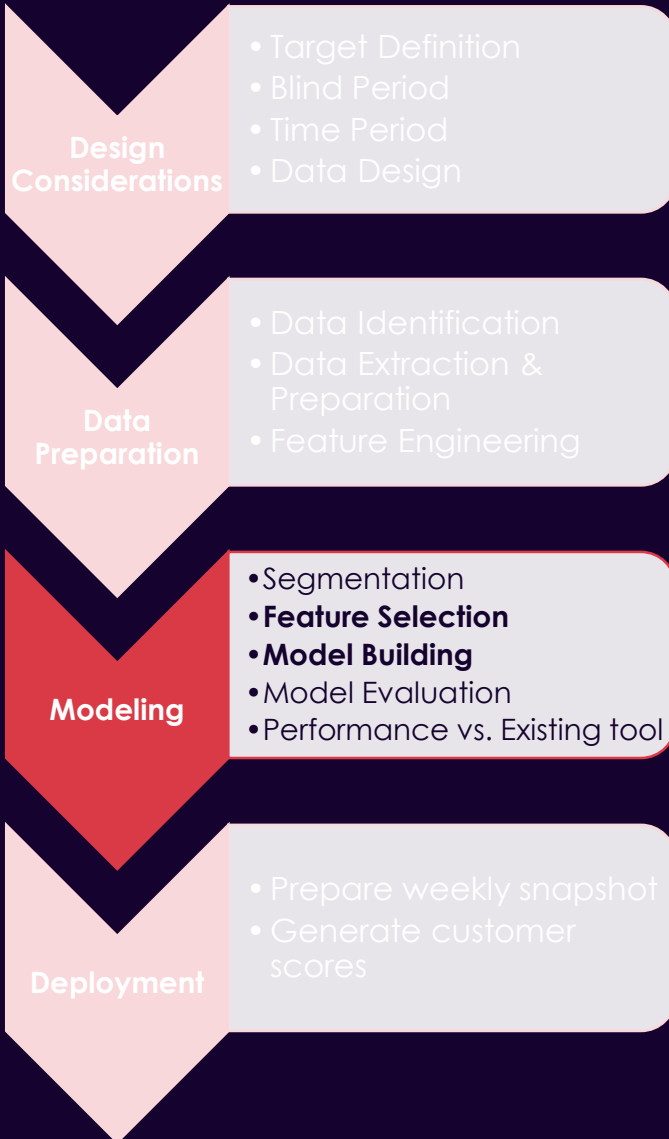


11 Models

- ✓ *SME / Corporate (4)*
- ✓ *Ditte individuali (3)*
- ✓ *Privati (4)*

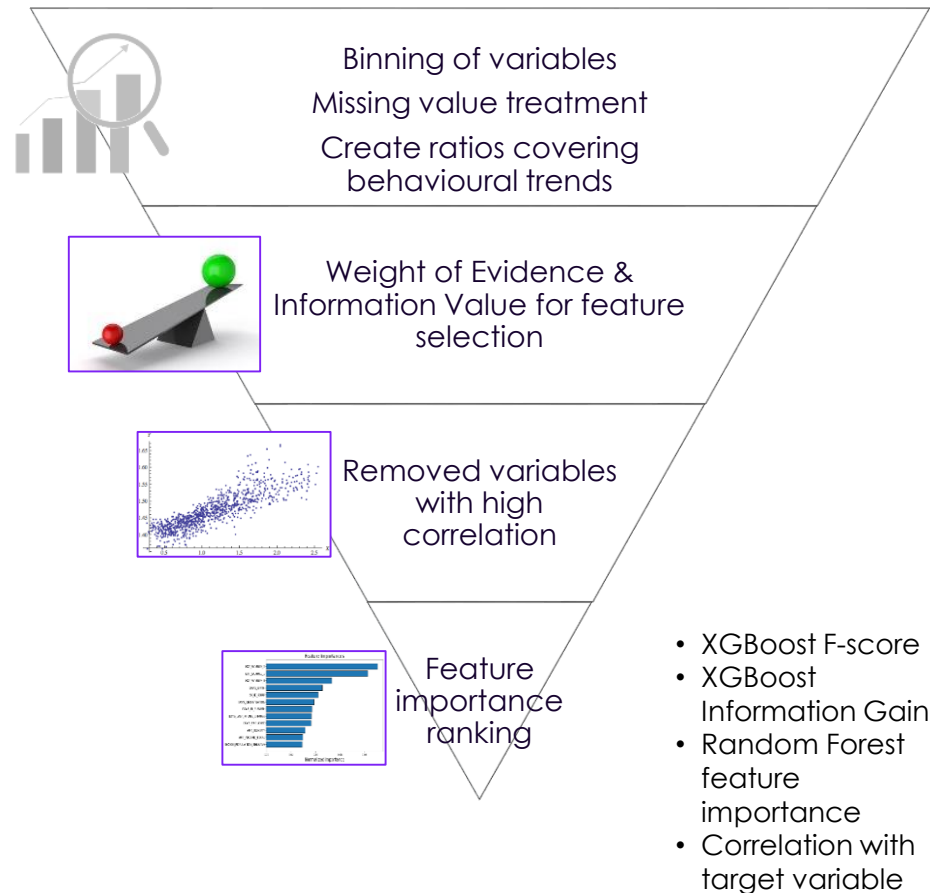
*Two types of credit – loans with installment (e.g. mortgage) & Line of credit (e.g. CA, overdraft)

* Exposure type depends on amount of exposure and availability of external risk data

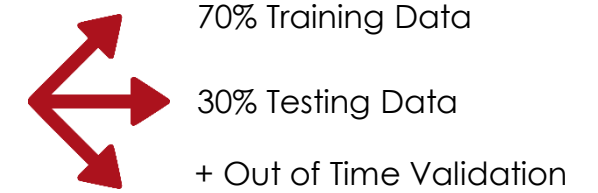


Feature Selection

A total of 3000+ features built; ~200 features selected for modeling stage



Model Building



Built Models for each segment identified



Approaches tried

- Logistic Regression
- Random Forest
- **XGBoost**



Model Choice

XGBoost: a complex tree-based ML algorithm that handles missing values internally

{centr/co:}



Model Evaluation

Model performance evaluated using the following key metrics:

- **AUC**
- **Precision**
- **Recall**

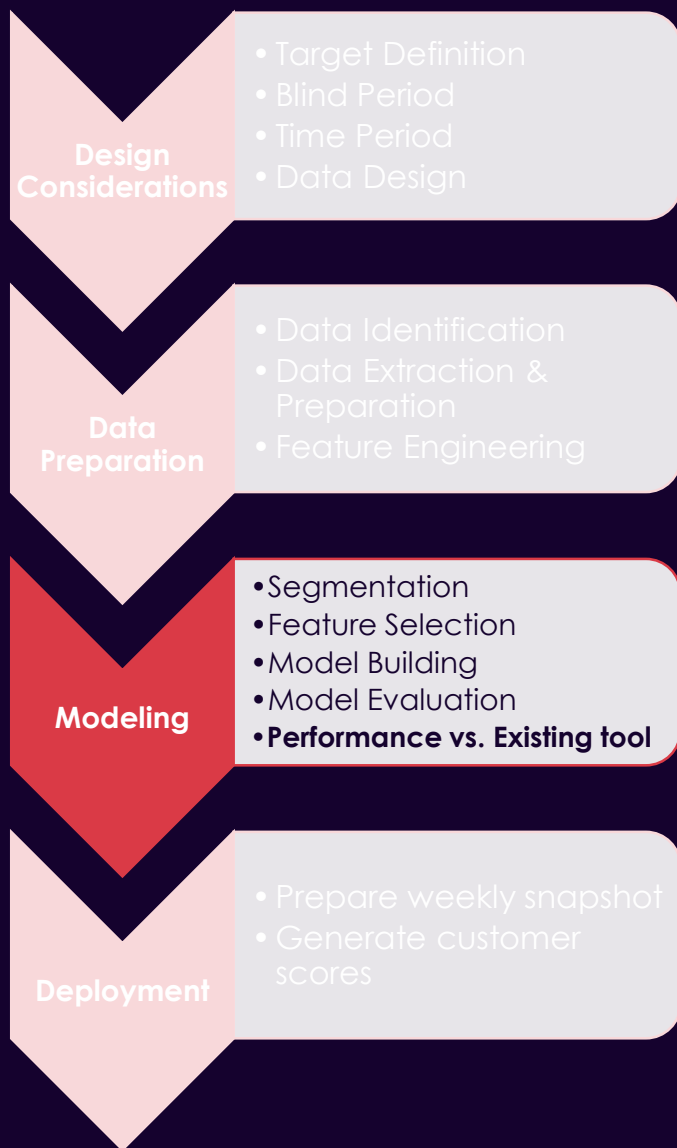


Select final models having:

- **Consistent** performance across Train, Test and Validation
- **Good** performance on various month (July 18-Oct 18) snapshots

Exposure Weights Combination	Learning Rate	Max Depth	Number of Boosters	Test AUC	Train AUC	10% Capture Rate Test	10% HitRate Test	10% Lift Test
[1, 1, 1]	0.1	2	50	89.0%	90.4%	63.7%	10.4%	6.4
[1, 2, 3]	0.1	2	50	88.7%	90.4%	62.0%	10.1%	6.2
[1, 1, 1]	0.1	2	100	90.4%	92.9%	64.4%	10.5%	6.4
[1, 1, 1]	0.1	3	50	89.9%	92.3%	64.7%	10.5%	6.5
[1, 5, 10]	0.1	2	50	87.7%	90.2%	59.6%	9.7%	6.0
[1, 10, 20]	0.1	2	50	87.4%	90.0%	59.9%	9.6%	5.9
[1, 2, 5]	0.1	2	50	87.4%	90.0%	61.6%	10.0%	6.2
[1, 2, 5]	0.1	2	100	90.3%	92.9%	63.4%	10.3%	6.3
[1, 2, 3]	0.1	3	50	89.7%	92.4%	65.1%	10.6%	6.5
[1, 2, 3]	0.1	2	100	90.2%	92.9%	64.4%	10.5%	6.4
[1, 20, 40]	0.1	2	50	87.0%	89.8%	56.2%	9.1%	5.6
[1, 5, 10]	0.1	2	100	90.0%	93.0%	62.0%	10.1%	6.2
[1, 10, 20]	0.1	3	50	89.0%	92.1%	63.0%	10.2%	6.3

centric



Data for 1 customer segment

Performance vs. Existing tool

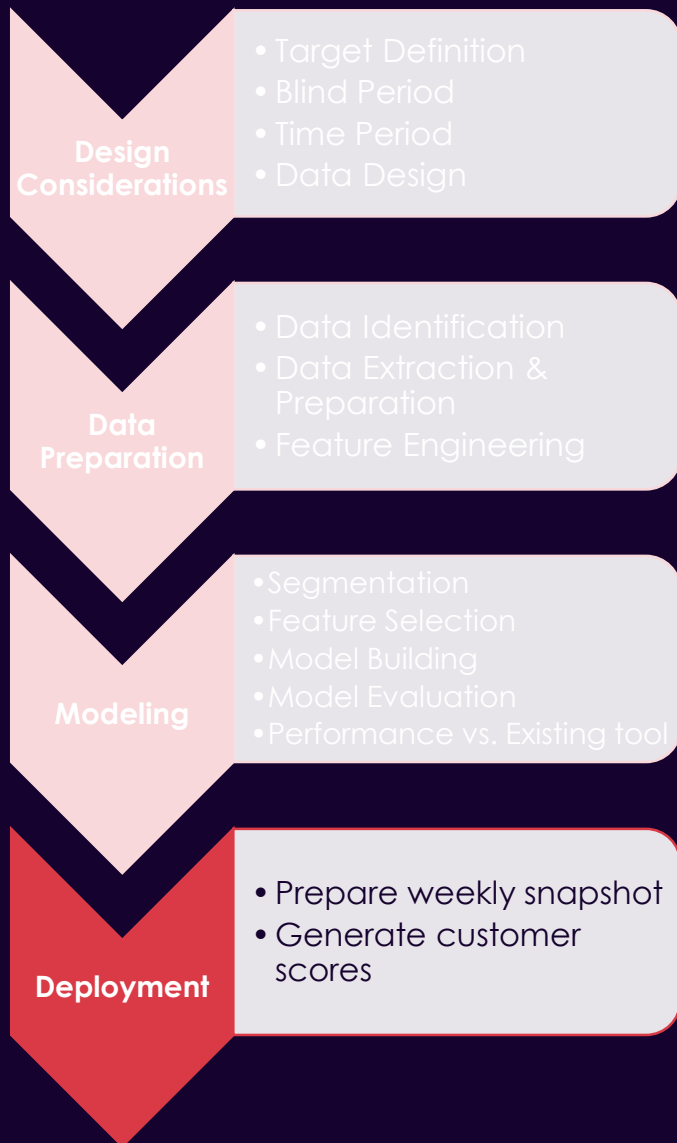
First Ever Events - Count of Customers Capture by Model vs CADR

Prediction	Model			CADR		
	Corporate	SME	Total	Corporate	SME	Total
Early	52%	57%	57%	24%	28%	28%
Late	24%	23%	23%	27%	31%	31%
Not Captured	24%	20%	20%	48%	41%	41%

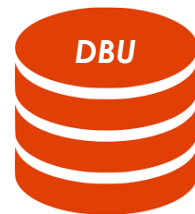
First Ever Events - Exposure Capture by Model vs CADR

Prediction	Model			CADR		
	Corporate	SME	Total	Corporate	SME	Total
Early	18%	58%	37%	8%	22%	14%
Late	38%	19%	29%	35%	36%	36%
Not Captured	44%	24%	34%	57%	43%	50%

Class	Hit Rate
Green	0.2%
Yellow	3.1%
Red	14.3%
Black	30.9%



✓ Weekly snapshot creation



Raw Data



Snapshot Generation

Deployment

- ✓ Weekly score generation
- ✓ **Model explanation (top indicators for each customer)**



on-premises



GPU based
AI infrastructure



Scheduled on a weekly base

{ centri/co :

Results & Lesson Learned

I risultati del nuovo modello CADR - EWI

Il nuovo modello è entrato in produzione nel 2020 per le aziende e per i privati.
Le performances rilevate sono molto incoraggianti e sono state apprezzate dal business:

- In grado di intercettare il 67% del rischio, in termini di esposizione, con un anticipo di 4-6 mesi rispetto al primo evento anomalo, con un effort pari a quello impiegato da CADR per intercettare il 33% del rischio
- Più efficiente del 54% rispetto alla CADR, in termini di teste, nell'intercettare preventivamente i clienti in default, comprendendo il 90% dei clienti effettivamente a rischio individuati dalla CADR.

La nuova CADR EWI è 2 volte più predittiva del modello precedente pur senza ampliare il perimetro di presidio.

E' in grado di intercettare la maggior parte delle posizioni a rischio default molti mesi prima che si manifestino le prime anomalie andamentali.

I risultati del nuovo modello CADR - EWI

	Solo primo evento di deterioramento						Tutti gli eventi					
	Per numero			Per esposizione			Per numero			Per esposizione		
	Rapporto EWI/CADR	Modello EWI	CADR	Rapporto EWI/CADR	Modello EWI	CADR	Rapporto EWI/CADR	Modello EWI	CADR	Rapporto EWI/CADR	Modello EWI	CADR
Intercettazioni Early Warning con preavviso 4-6 mesi	1.54	33%	21%	2.03	67%	33%	1.21	33%	27%	1.54	62%	40%
Intercettazioni Early Warning con preavviso 3-6 mesi	2.51	72%	29%	1.49	77%	51%	2.04	68%	33%	1.33	69%	52%
Intercettazioni complessive con preavviso 1-6 mesi	1.83	89%	49%	1.17	90%	77%	1.60	83%	52%	1.14	79%	69%

Conseguenze del nuovo modello CADR - EWI

L'efficacia del nuovo modello Cadr rispetto al precedente ha comportato modifiche nella gestione del credito problematico:

- ✓ Con la vecchia Cadr il gestore della posizione era facilitato nella definizione della strategia. Attendeva in modo passivo la classificazione della posizione in cadr rossa o nera. Dopodichè l'intervento del gestore era molto rapido e prevalentemente indirizzato al tentativo di sistemazione della anomalia andamentale conclamata che aveva determinato il pessimo colore cadr.
- ✓ Con la nuova Cadr il gestore non è spesso in grado di capire le cause della classificazione. Le posizioni vengono ora classificate in cadr nera o rossa ancor prima che si vengano a determinare le prime anomalie andamentali .
Il gestore della posizione si trova così di fronte un cliente ancora sano che però sta adottando comportamenti impercettibili che statisticamente lo condurranno al default.
Il gestore della posizione deve quindi effettuare una verifica approfondata a 360° del cliente, dall'analisi bilanci alla centrale rischi al colloquio col debitore, e valutare misure cautelative quali l'avvio del progressivo disimpegno o l'acquisizione di nuove garanzie

Next steps

Sviluppi futuri Nuova Cadr

- 1. Potenziamento delle performance del modello di early warning tramite la crescente introduzione di banche dati esterne*
- 2. Adozione di strumenti di gestione del rischio di credito evoluti (Peg) per sfruttare al meglio le previsioni EWI*



fabio.polto@sella.it

stefano.priola@centrico.tech

federico.angaramo@centrico.tech

Centrico S.p.A

sede legale in Piazza Gaudenzio Sella n. 1, 13900 Biella (BI)

Iscritta al Registro di Biella e Vercelli, codice fiscale 02668670025

Società appartenente al Gruppo IVA Maurizio Sella S.A.A. con partita IVA 02675650028

Soggetta all'attività di direzione e coordinamento di Banca Sella Holding S.p.A.

PEC: centrico@pec.centrico.tech MAIL: info@centrico.tech

Banca
Sella

{ centric^o :