AI4Business
Developing AI tools

Roadmap AI4Business

Part I	Part II	Part III	Part IV	Part V
Introduction to AI	Developing AI tools	Data and Value	Deploying AI	Monitoring

Table of contents

1 Typical AI team

AI team members

Core team

- Data Science Team Lead
- Data Scientist
- Data Engineer

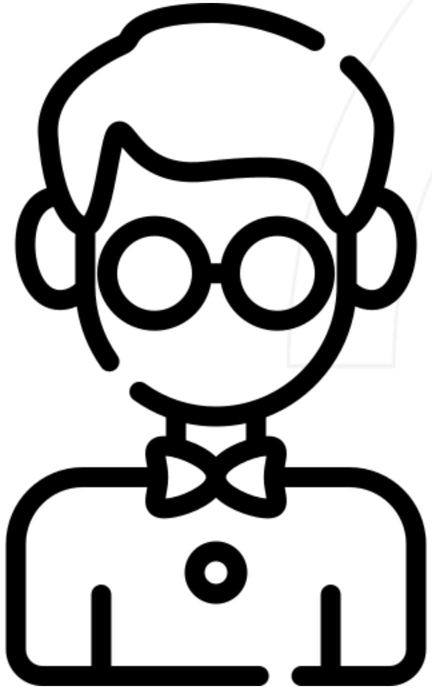
Supporting roles

- Business Intelligence Analyst
- Database Administrator
- DevOps / MLOps

Data Science Team Lead

Role:
Manages the data science team and plan projects

Goal:
Make sure business goal are aligned with data vision to create value



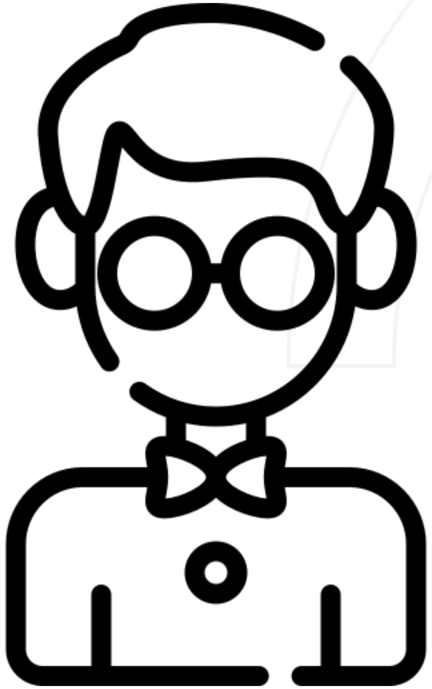
Toolkit:
SQL, Python

Skills:
Leadership and project management
Interpersonal communication
Data analysis + visualization
Predictive modeling
Business mindset

Data Scientist

Role:
Collects, analyzes and interprets the data to propose business decisions

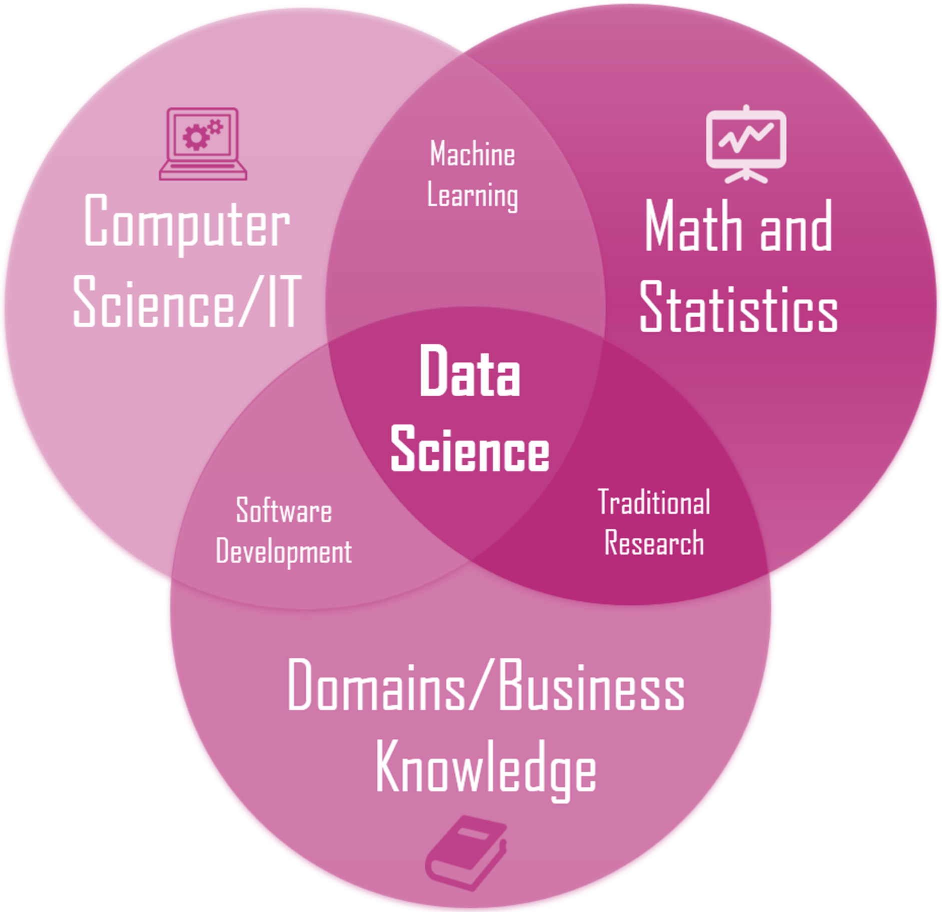
Goal:
Construct the means for extracting business-focused insights from data



Toolkit:
Python, Scikit-learn, Pandas, Numpy, Tensorflow, SQL

Skills:
Predictive modeling
Data visualization and storytelling
Maths and Stats
ML and DL algorithms

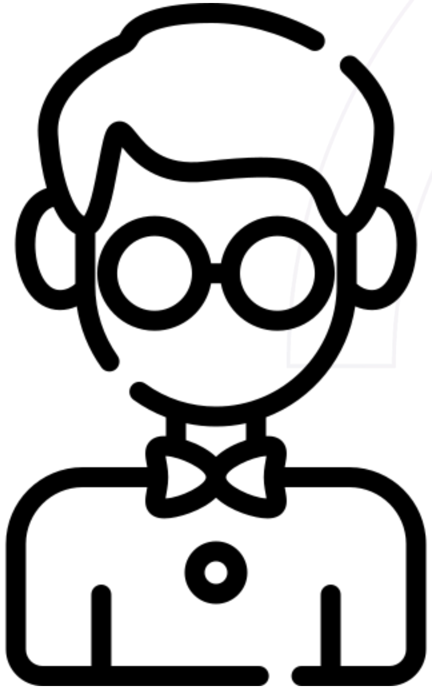
Data Scientist



Data Engineer

Role:
Design, build and implement the data infrastructure that fuels machine learning and AI analytics

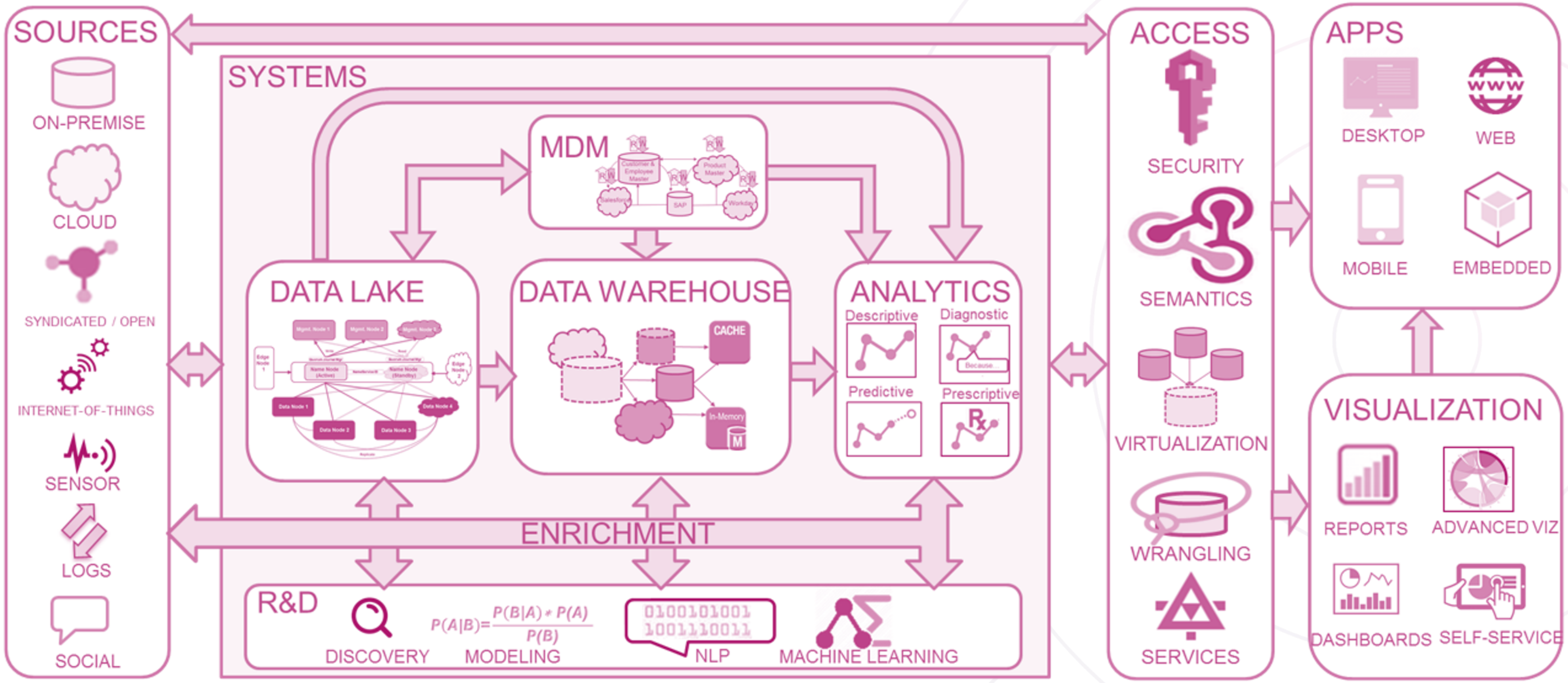
Goal:
Integrate and manage data from various sources to support data scientists with systems to build predictive pipelines



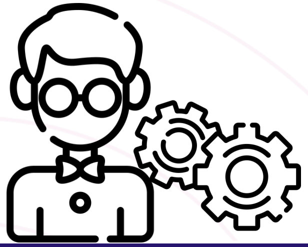
Toolkit:
Python, SQL, ETL tools

Skills:
Data warehousing solutions
Database systems
Data modeling
Data APIs

Data Engineer

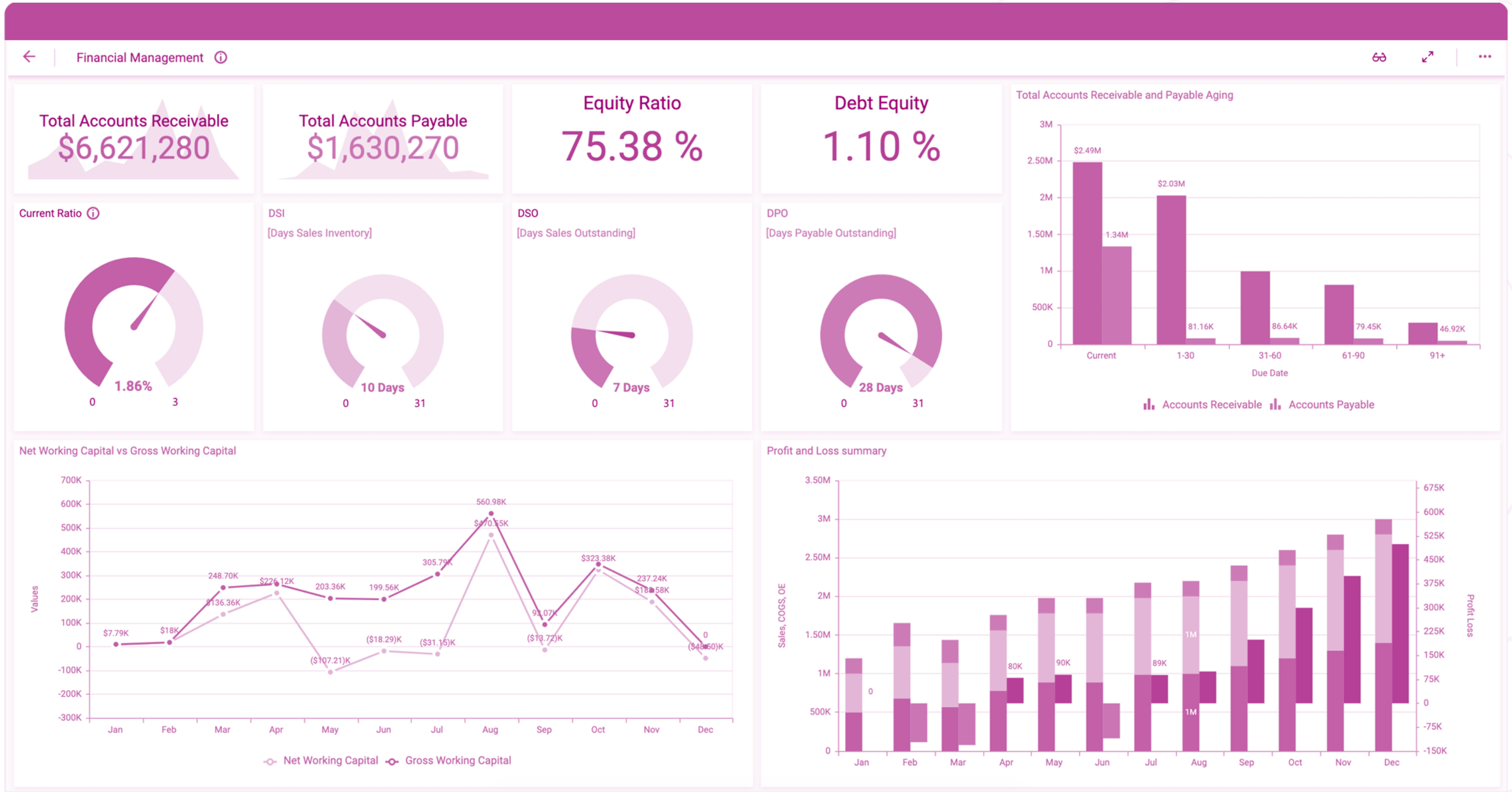


Supporting roles



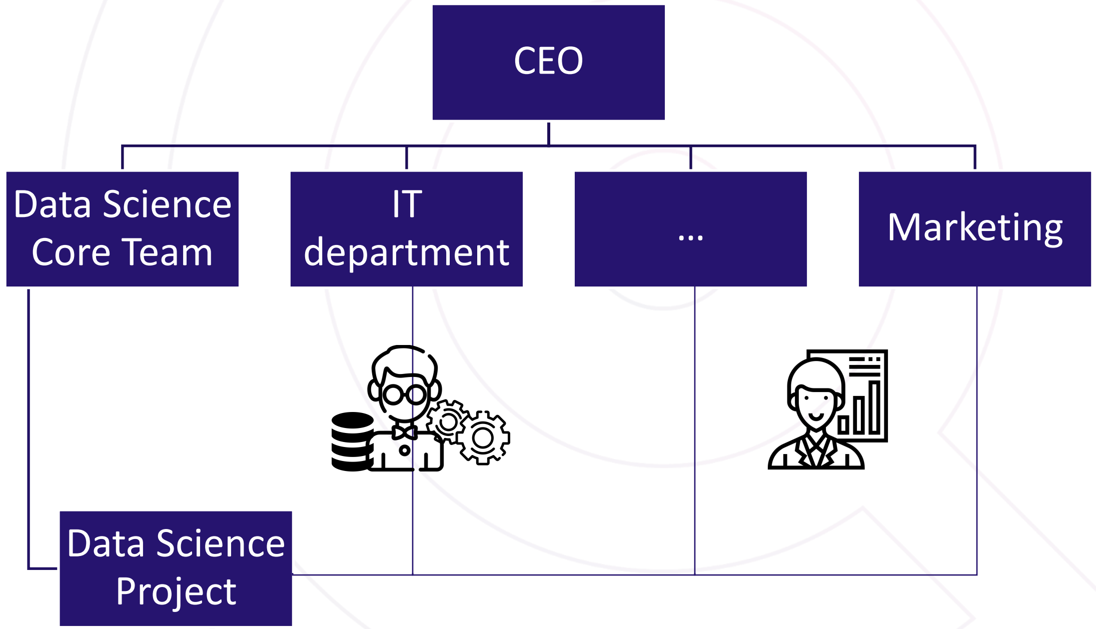
	Business Intelligence Analyst	Database Administrator	DevOps / MLOps
Role	Improves business processes via data; intermediary between DS team and department	Ensures data availability to a DS team	Integrates ML solutions into existing system
Skills	Data modeling Data visualization tools Communication skills Business understanding	Backup & Recovery Data modeling and design Distributed Computing Data security	Cloud Infrastructure Virtualization Knowledge of CI/CD systems
Department	Every department	IT Department	IT Department

Business Intelligence Analyst



Organizational structure

- Data Science Team is
  - a separate organizational unit
  - reporting directly to CEO
- One team serves the whole organization in a variety of projects
- Supported by:
  - Database Administrators + MLOps from IT
  - BI Analysts at a departmental level
- Data Science Team should participate in activities of business units



Exercise

Who said what?

- My dashboard shows that sales are down every April in the past 4 years
- My model predicts that sales in November will go up
- I created a way for you to access sales data on a monthly level
- You are assigned to this sales project based on your retail experience
- Your ML model will be up and running in the application by tomorrow

Role

- …
- …
- …
- …
- …

Exercise

Who said what?

- My dashboard shows that sales are down every April in the past 4 years
- My model predicts that sales in November will go up
- I created a way for you to access sales data on a monthly level
- You are assigned to this sales project based on your retail experience
- Your ML model will be up and running in the application by tomorrow

Role

- BI analyst
- Data scientist
- Data engineer
- Team lead
- MLOps engineer

2 Project stages

Three lenses of due diligence

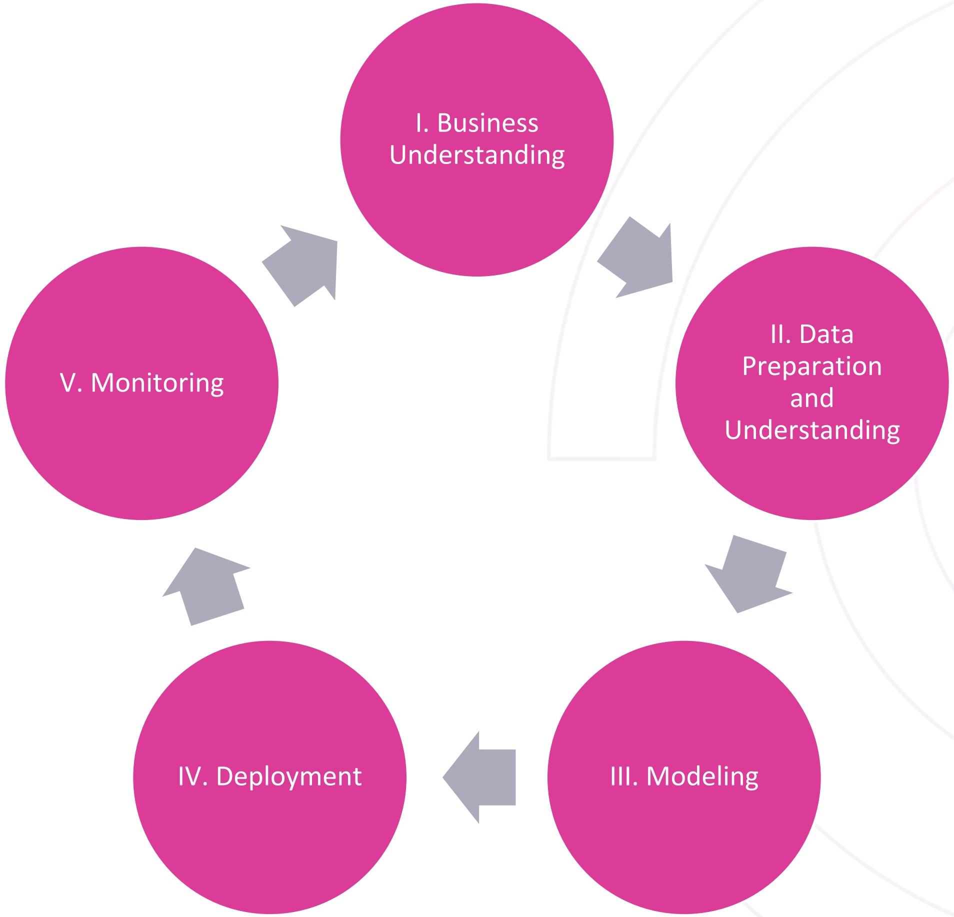
- Due diligence before the start of any project:

- Business viability: positive value creation
  - increase revenue, lower cost, boost efficiency or launch new business

- Technical feasibility: AI system can be built
  - meet desired performance, data availability, engineering timeline, etc.

- Human desirability: project is really wanted and ethically ok
  - AI developed/used by people

Data Science Life Cycle



I. Business Understanding

- Identify an opportunity to create value
  - Which part of your company process workflow can benefit from AI?
- How can AI help us?
  - Automate existing processes and facilitate human-machine collaboration
  - Improve existing algorithms to become more accurate or reliable
  - New business opportunities
- Pin down the project’s goals
  - How is AI going to solve the problem?

II. Data Preparation

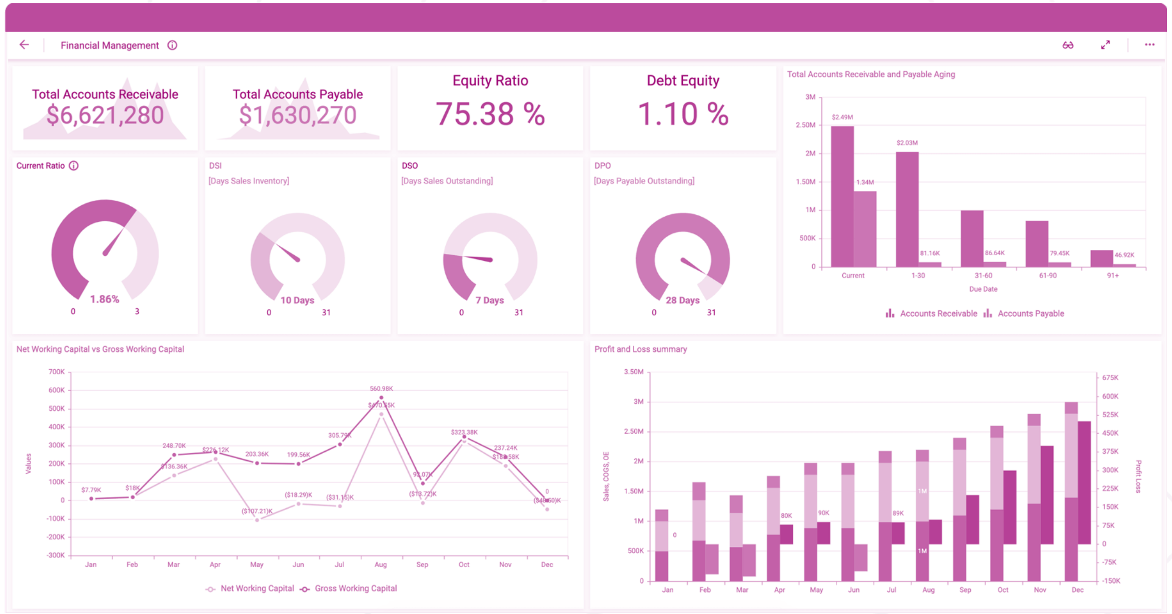
- Data collection
  - Internal & external data sources to acquire relevant and comprehensive data
  - Focus on data quality

- Data processing
  - Cleaning to deal with missing and inconsistent data
  - Preparation for modeling phase
  - Can be very time consuming

- Need a structured way to deal with data and centralize data flow

II. Data Understanding

- Summarize the main characteristics of the data set
- Represent dataset visually in a dashboard
- Understanding the patterns and bias in the data
- Gain insight into the data
- Assesses quality of the data



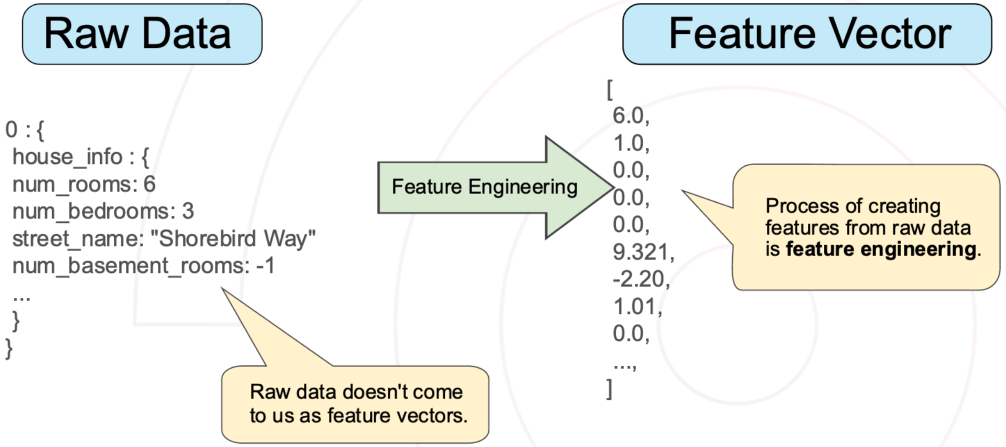
Key questions regarding data

- What kind of data do I have available?
- Where can I find it in my organization?
- Who owns the data and am I allowed to use it?
- What is the format/quality of the data?
- Can I trust my data?
- Possible to enrich own data with external data?

- Is this the right data to solve my business problem at hand?

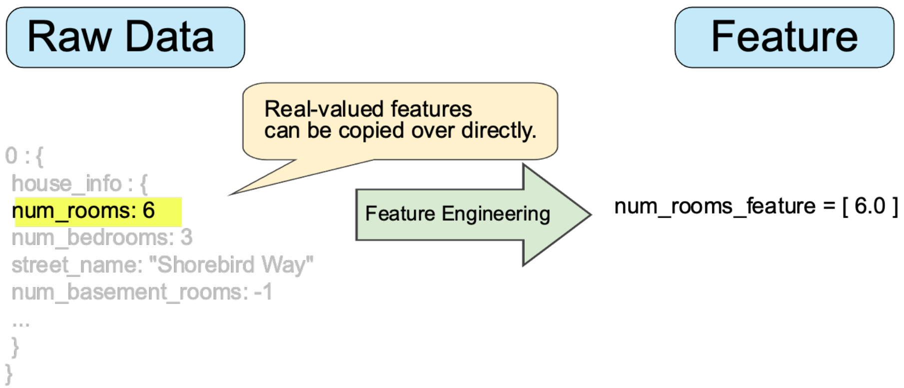
III. Modeling – Feature Engineering

- Transform raw data into usable features:
  - [Google – ML Crash Course]

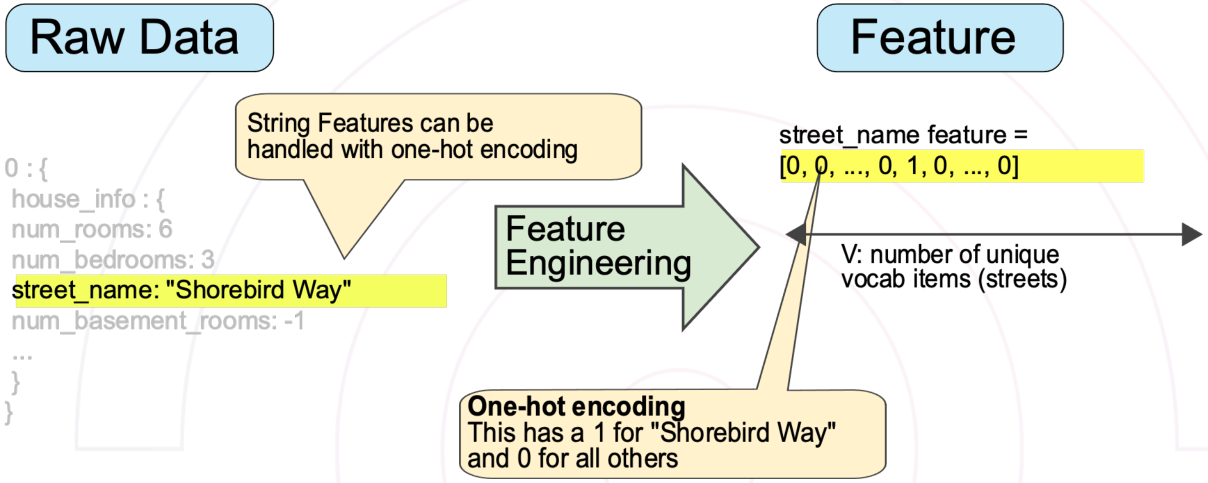


- Feature construction
  - creating new features from the ones that you already have
- Feature selection
  - Remove irrelevant features that add more noise than information

III. Modeling- Feature Engineering



Quantitative or numerical features

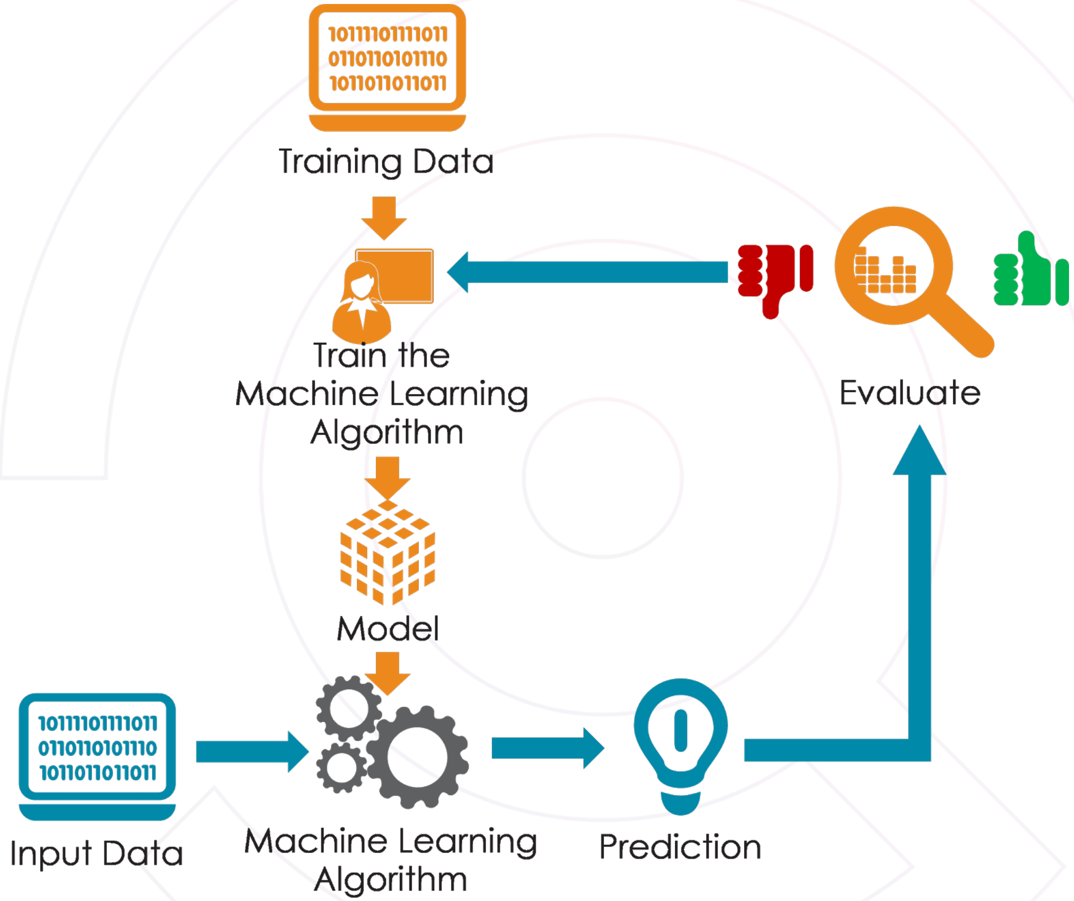


Qualitative or categorical features

Street	Feature vector
Charleston Road	[1,0,0,0]
North Shoreline Boulevard	[0,1,0,0]
Shorebird Way	[0,0,1,0]
Rengstorff Avenue	[0,0,0,1]

III. Modeling – Machine Learning

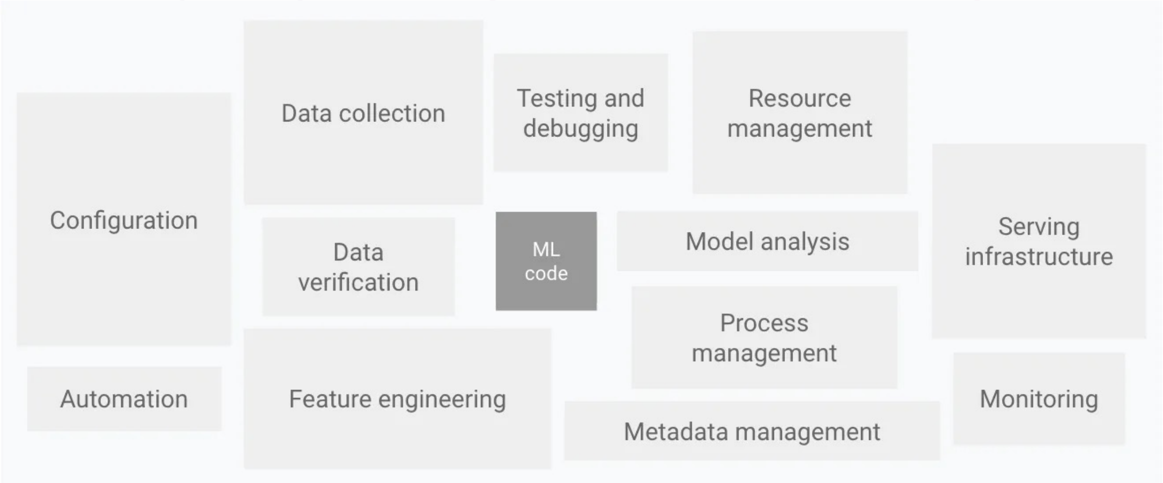
- Train various algorithms to develop models
- Evaluate model performance on new unseen data samples
- Ensure that the outcomes make sense and are significant
- Typically iterative process
    - Review model
    - Review data



IV. Deployment

- Integration of the ML model into an existing production environment
  - ML code only constitutes a tiny part of the full production architecture

- Done by MLOps engineer



[Google Cloud - MLOps]

- Proof of Concept (PoC) to production gap
  - Many researched ML solutions don’t see daylight

V. Monitoring

- The real world is constantly changing
  - Data drift: the distribution of input features changes (e.g., houses become smaller over time because of space scarceness)
  - Concept drift: the mapping from features to target changes (e.g., popularity for small houses makes these more expensive)

- Ensure that algorithms keep doing a good job once deployed
  - Constantly evaluate their performance with regard to a baseline
  - Identify degrading solutions early on

- Brainstorm statistics/metrics to track over time
  - Visualize in a dashboard
  - Set thresholds for alarms
  - Iterative process: adjust metrics + thresholds over time

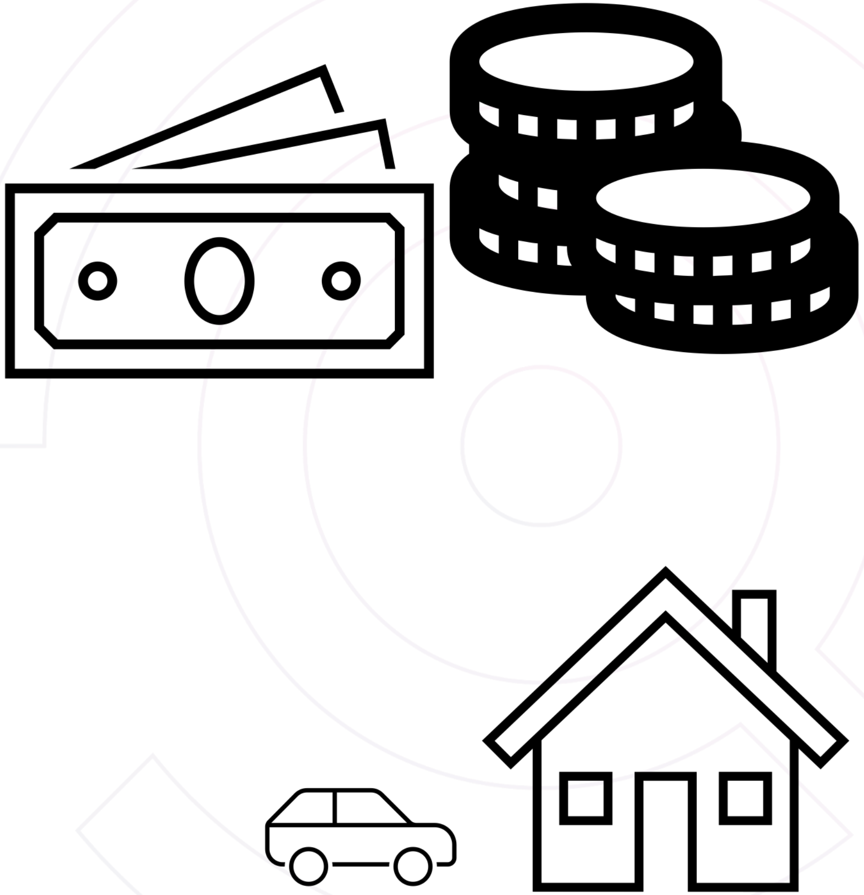
3 The big picture

Problem definition

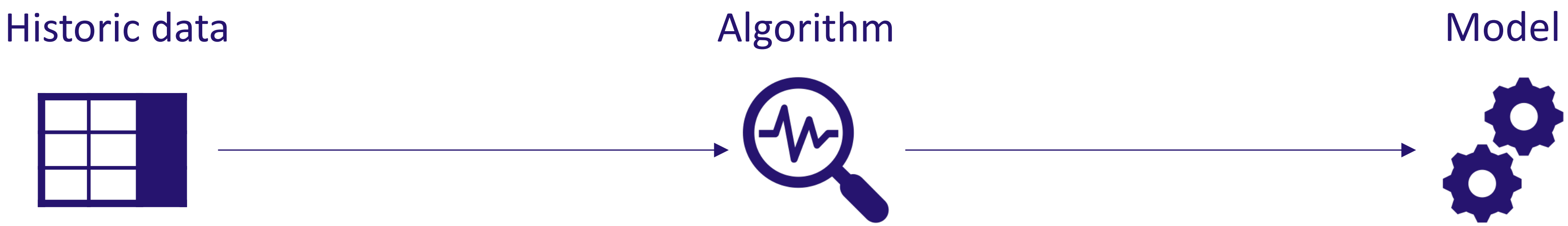
- Automatic feeding system for chickens
- Deciding how much each chicken should get?
- Problem
  - ✓Roosters and hens look almost the same
  - ✓Roosters are bigger and require more nutrition

Problem definition

- Default prediction for bank loans
- Who will be able to repay the loan amount?
- Problem
    - ✓ Default causes big losses for the bank
    - ✓ Find good profiles to grant a loan and avoid risky profiles



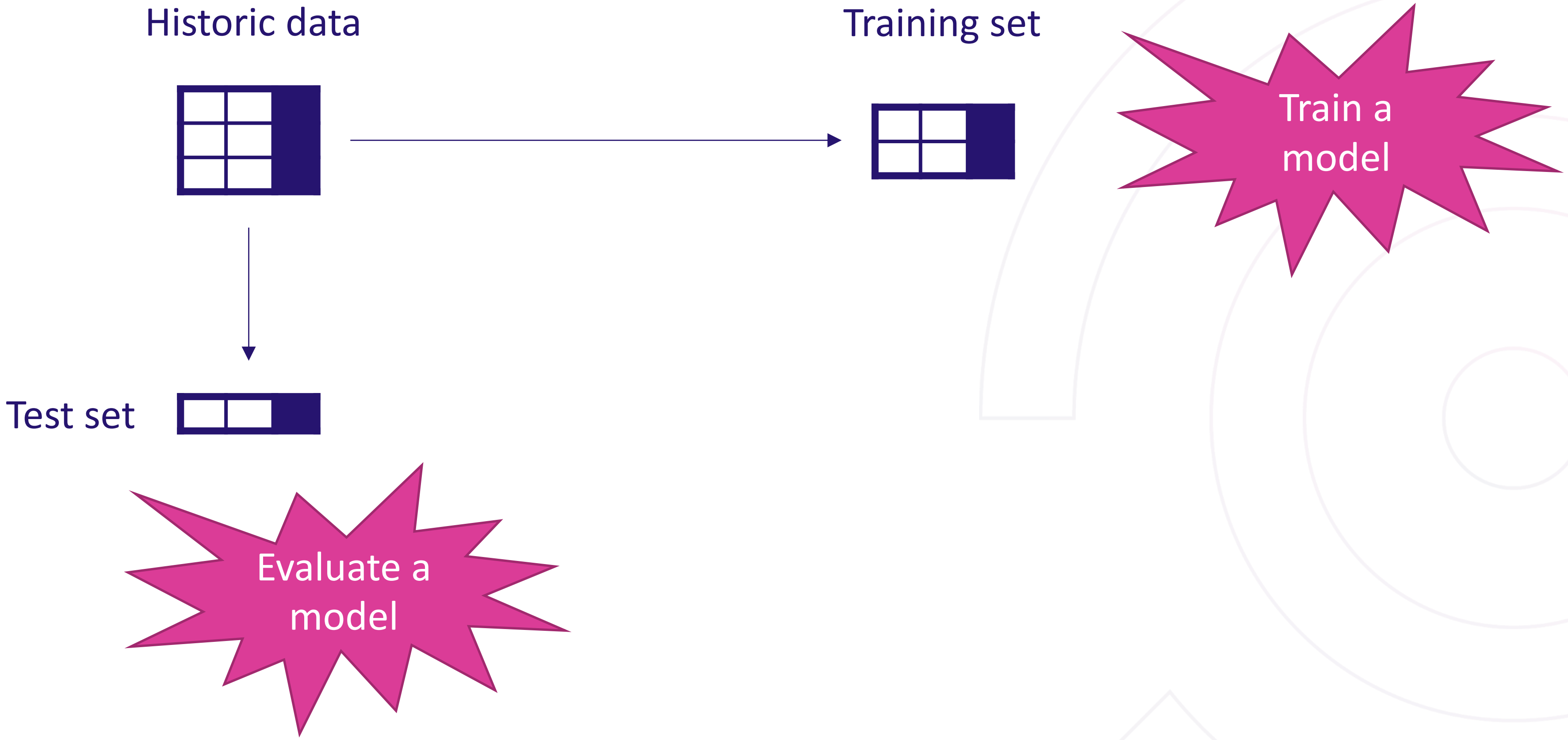
From data to model



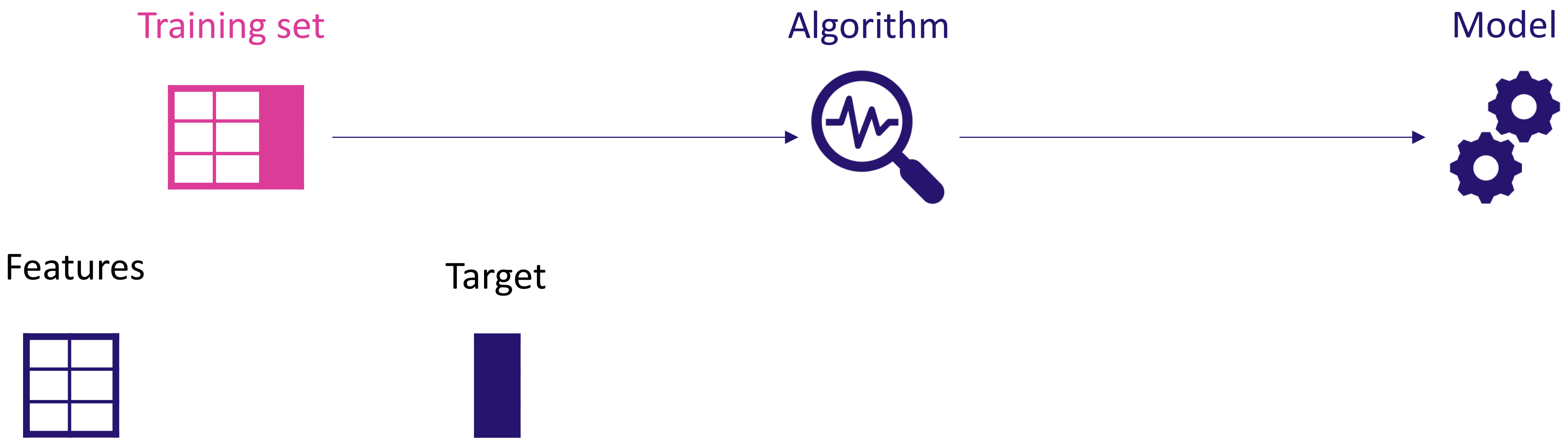
Complete dataset

Historic data

Data partitioning



Training data



Unstructured data

Training set

Algorithm

Model

Features

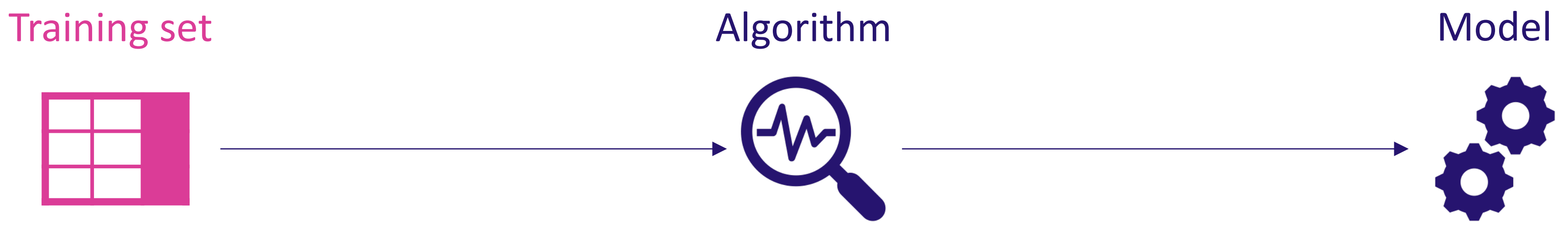
Target

Hen

Rooster

Images

Structured data



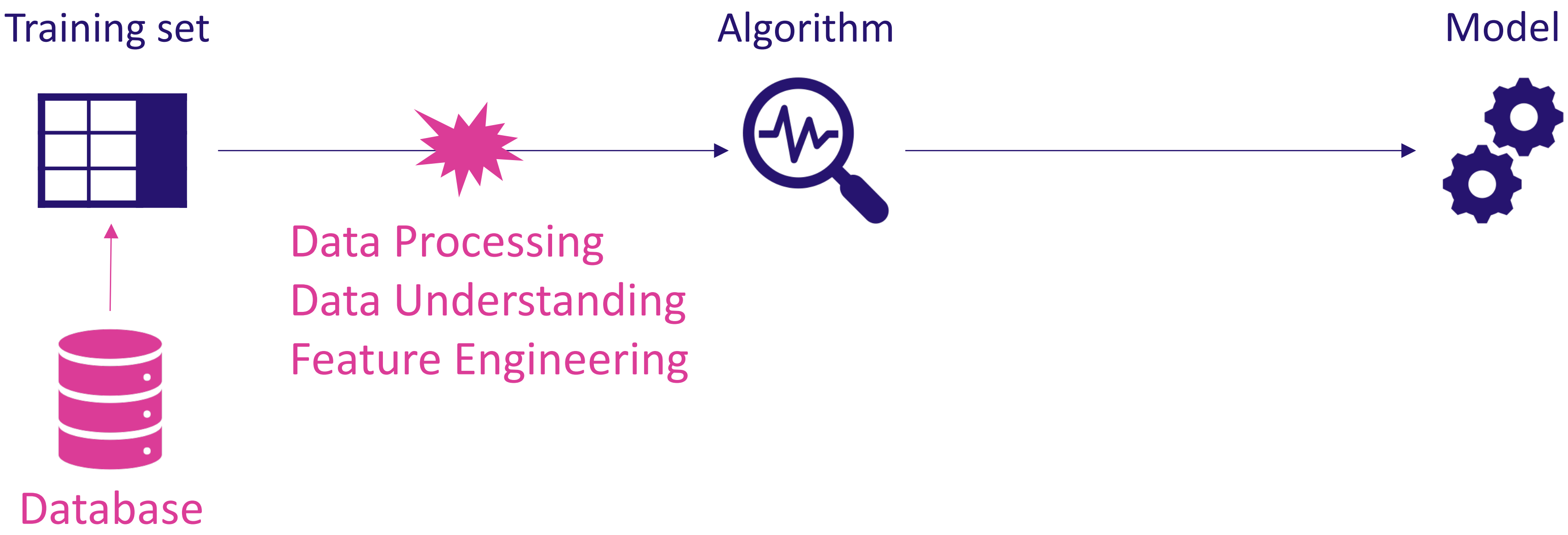
Customer Name	Age	Education	Gender	Default
Joris Maes	62	Low	Male	Yes
Pieter Claes	32	High	Male	No
Nina Peters	25	Low	Female	No

Features

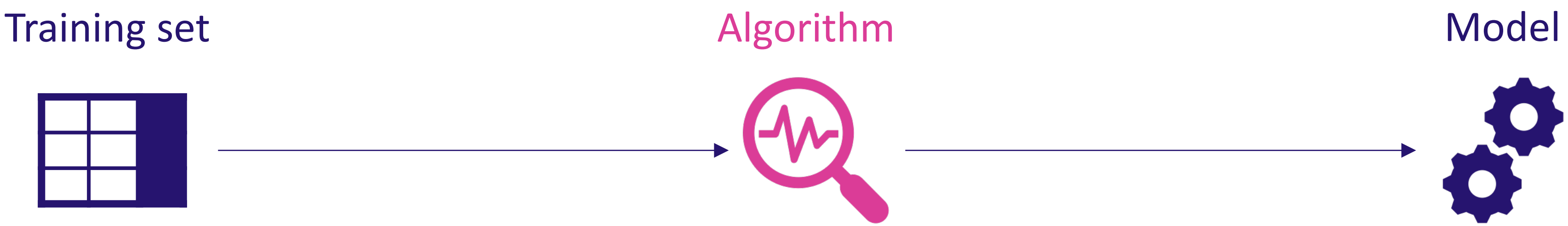
Target

Table

Data preparation



Machine learning

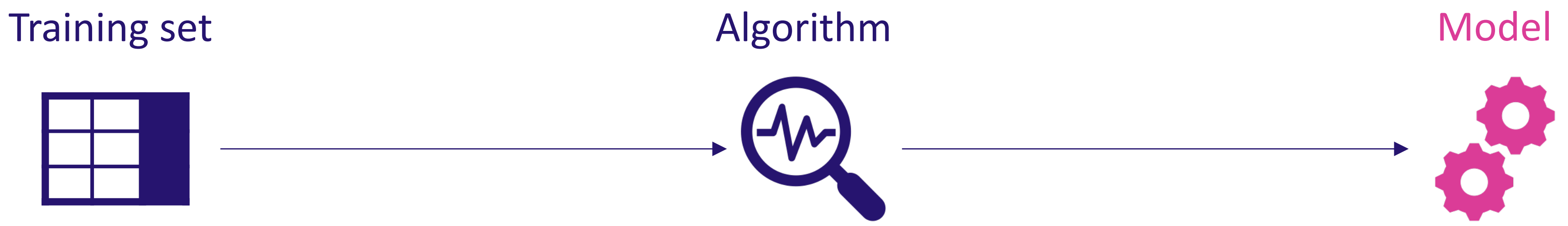


Look for patterns that map features to the target

Customer Name	Age	Education	Gender	Default
Joris Maes	62	Low	Male	Yes
Pieter Claes	32	High	Male	No
Nina Peters	25	Low	Female	No

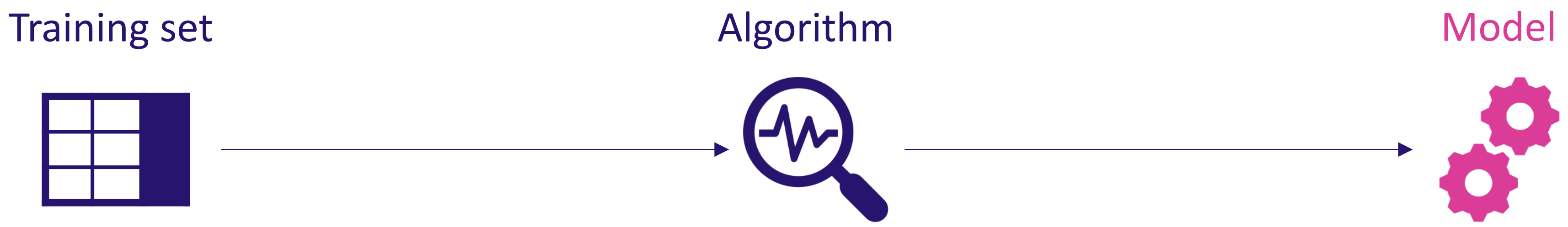
Features | Target

Model predictions



Customer Name	Age	Education	Gender	Default
Joris Maes	62	Low	Male	Yes
Pieter Claes	32	High	Male	No
Nina Peters	25	Low	Female	No
Erica Maes	27	Low	Female	???

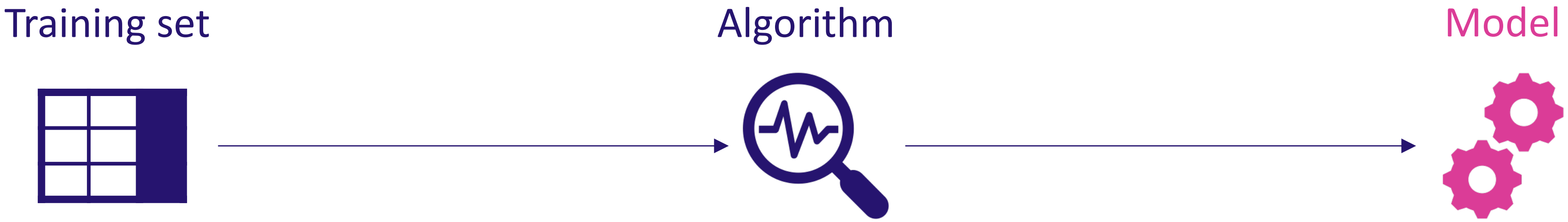
Model candidate #1



Customer Name	Age	Education	Gender	Default
Joris Maes	62	Low	Male	Yes
Pieter Claes	32	High	Male	No
Nina Peters	25	Low	Female	No
Erica Maes	27	Low	Female	No

IF Customer Young **&** Female
Then Default No

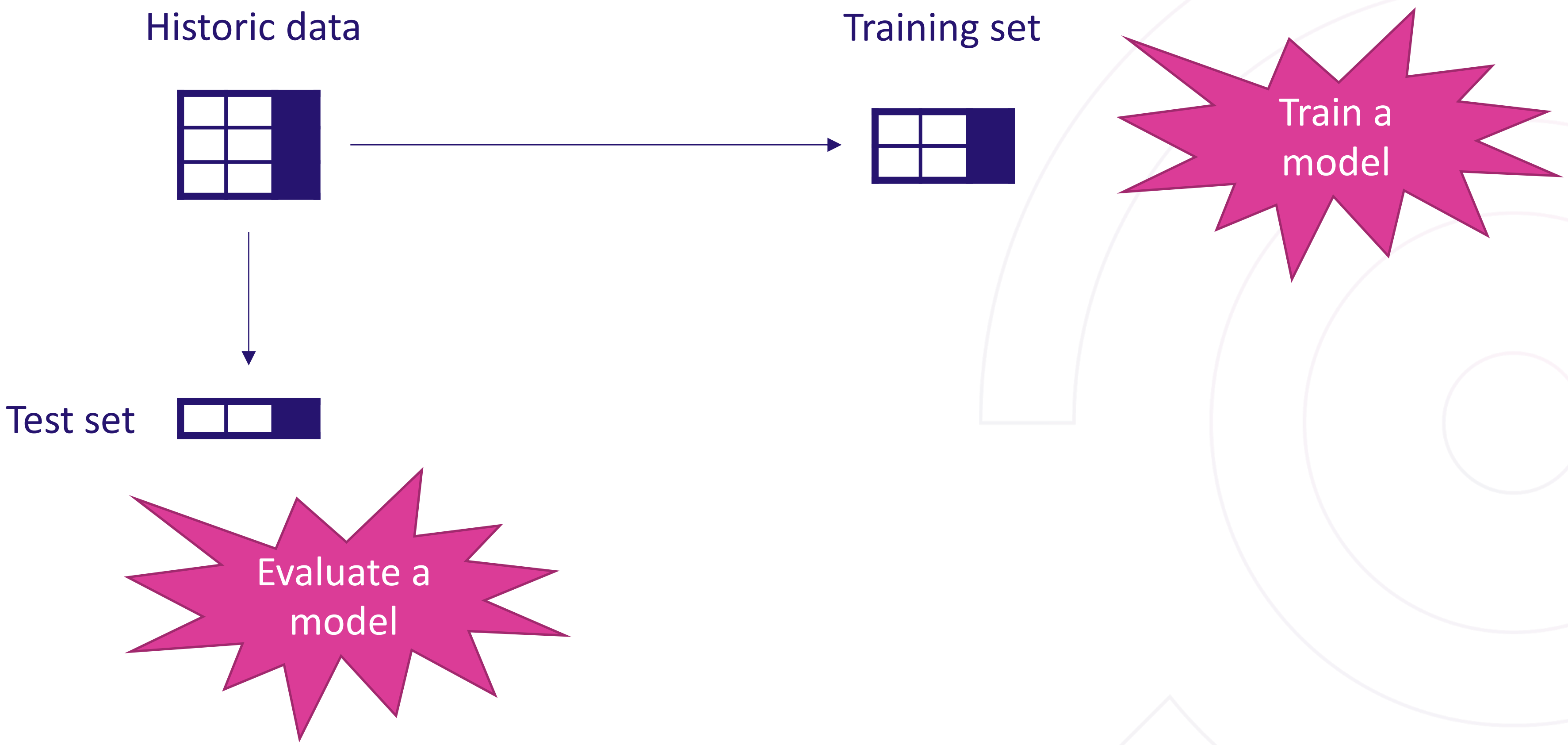
Model candidate #2



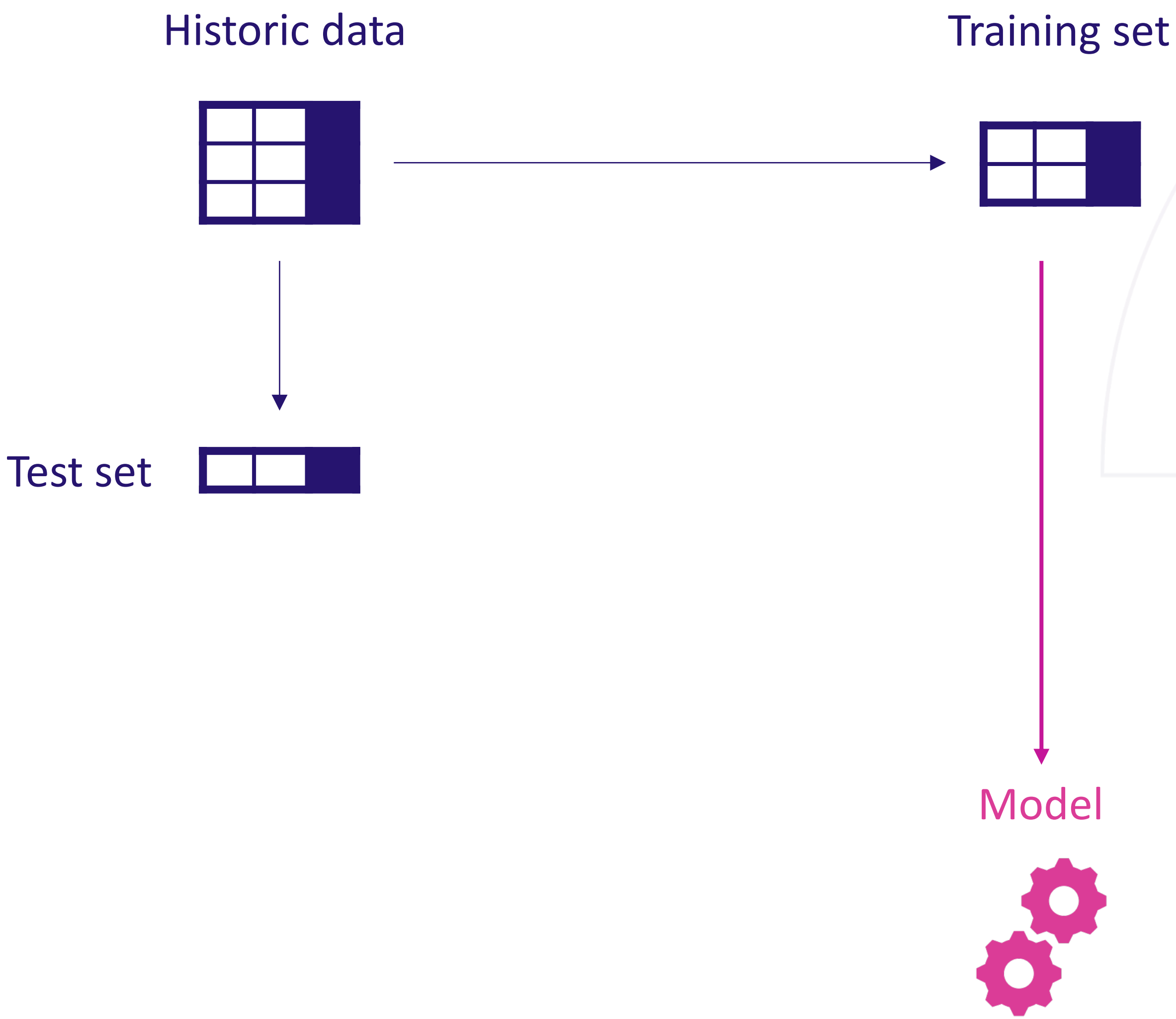
Customer Name	Age	Education	Gender	Default
Joris Maes	62	Low	Male	Yes
Pieter Claes	32	High	Male	No
Nina Peters	25	Low	Female	No
Erica Maes	27	Low	Female	Yes

IF Customer Name “Maes”
Then Default Yes

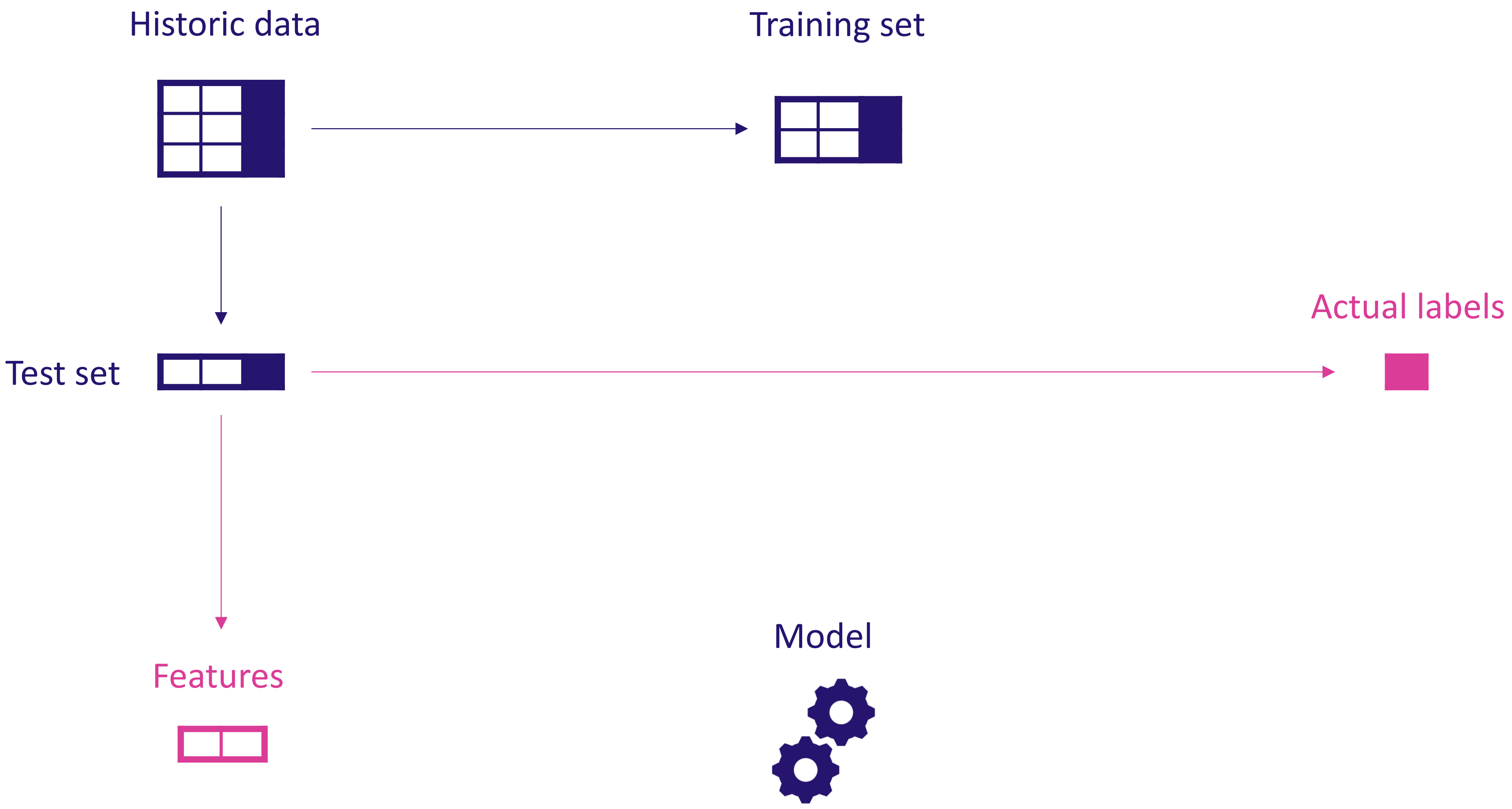
Which model to choose?



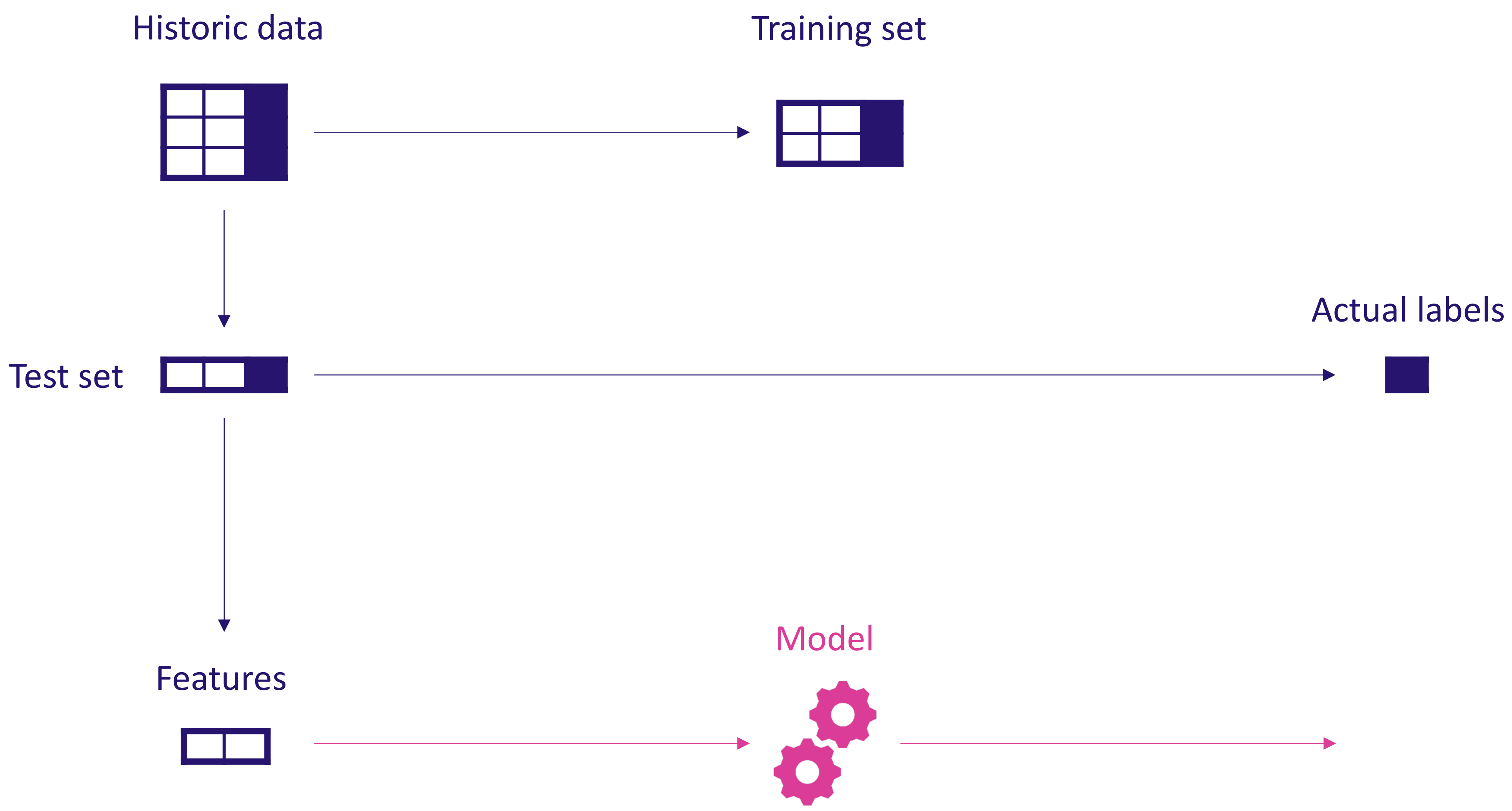
Trained model



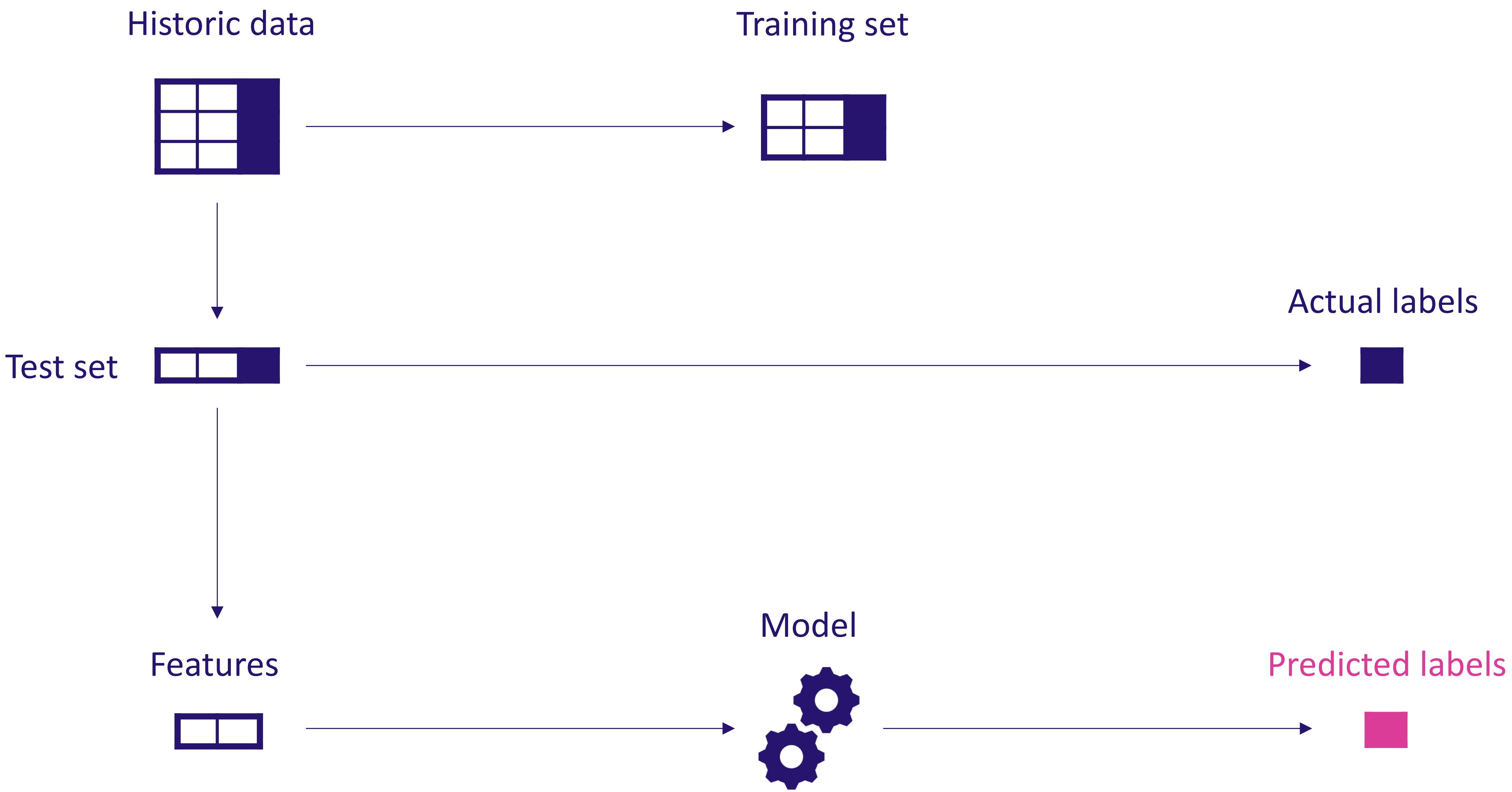
Evaluating a Model



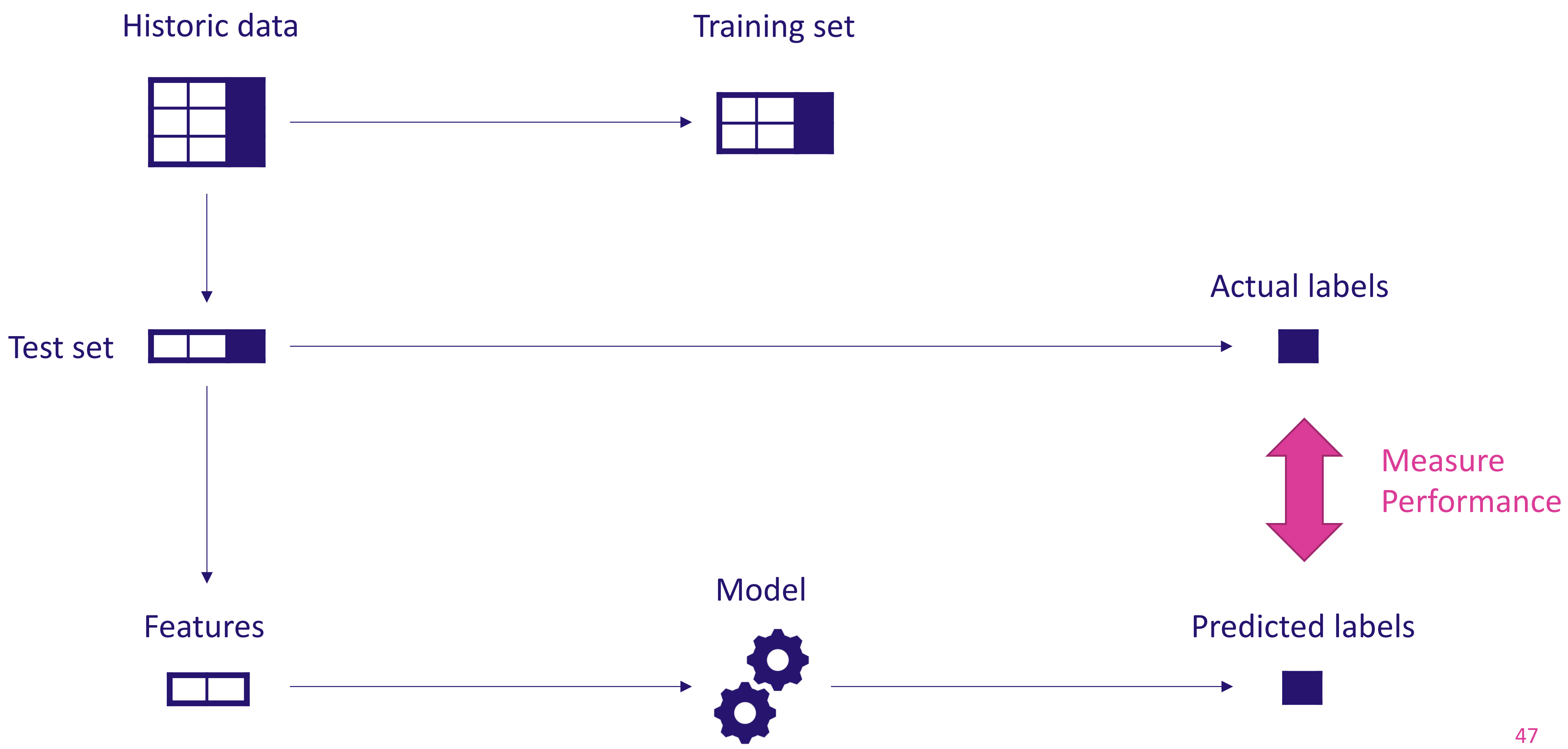
Evaluating a Model



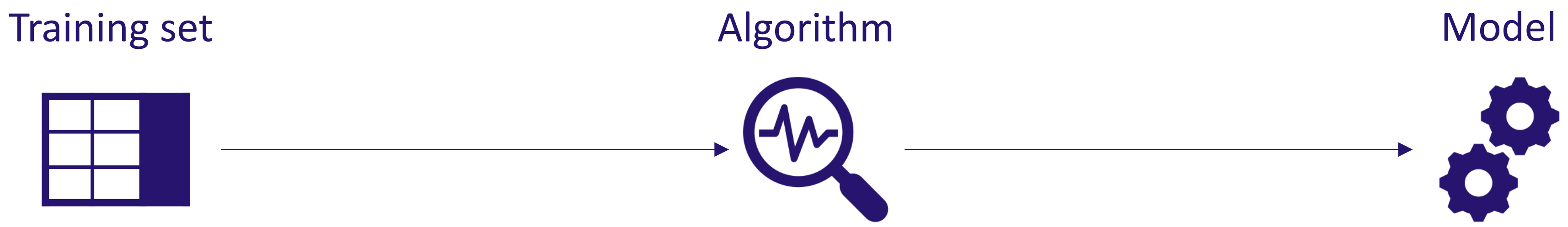
Evaluating a Model



Evaluating a Model

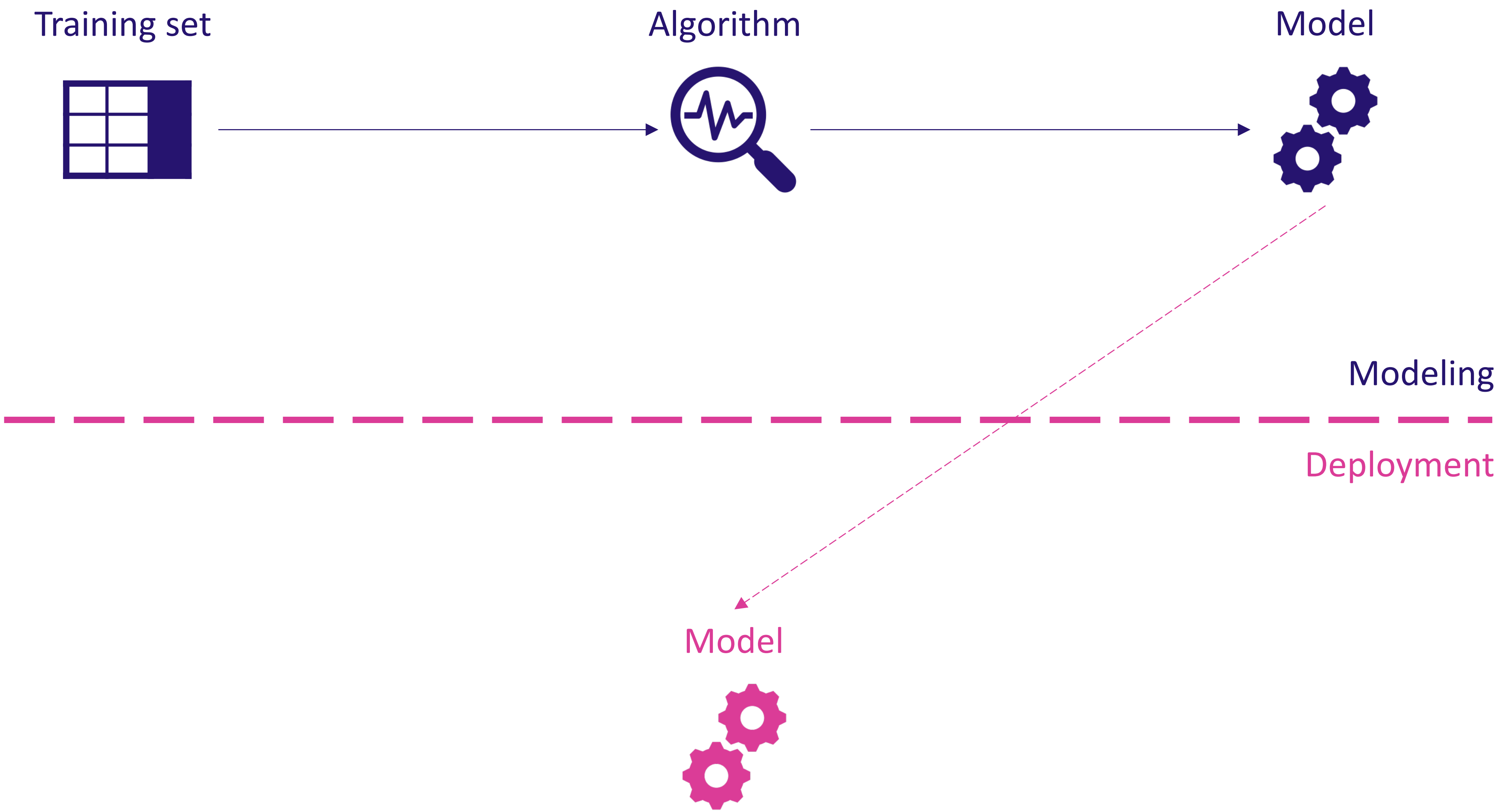


Modeling

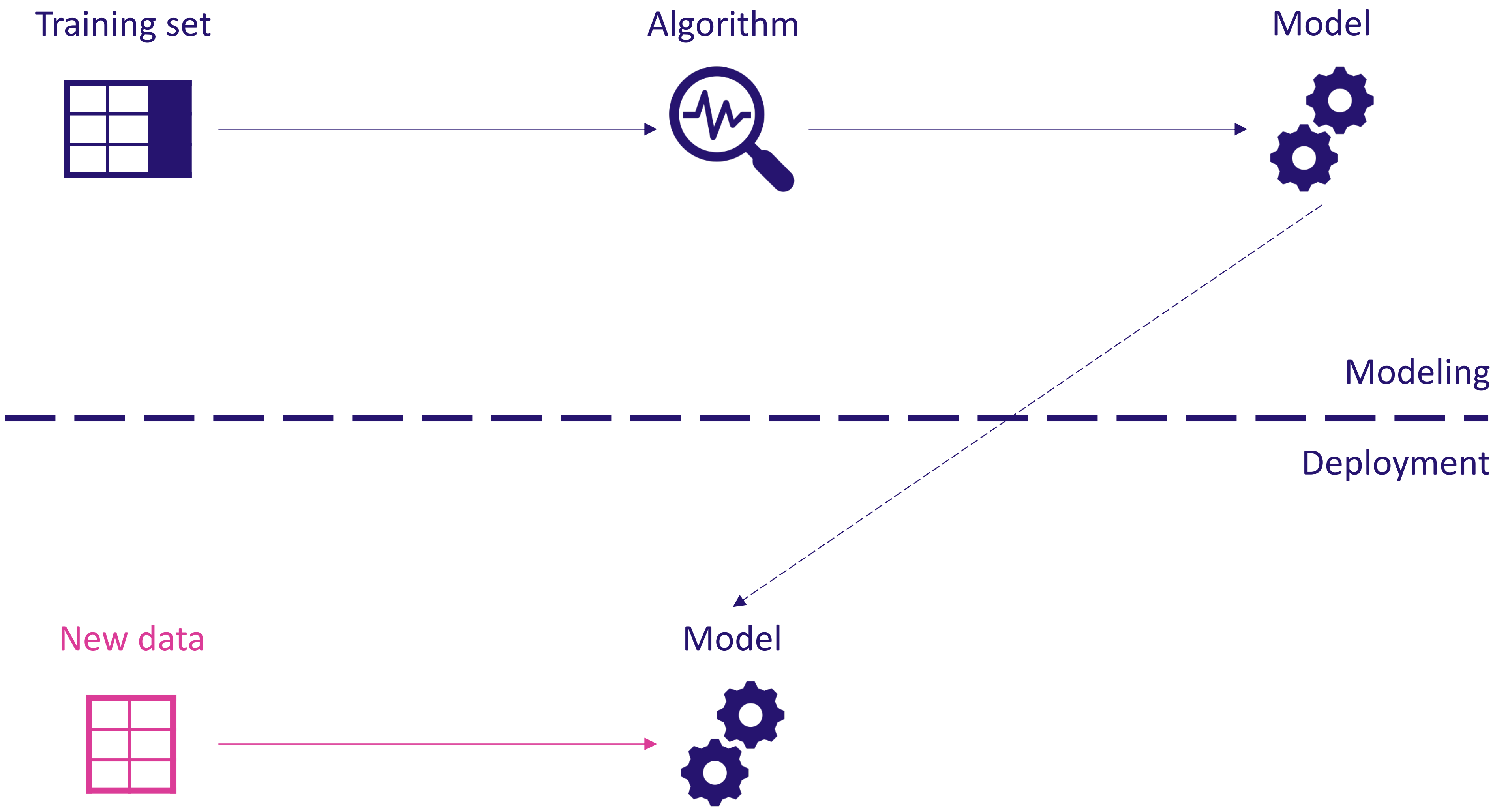


Modeling

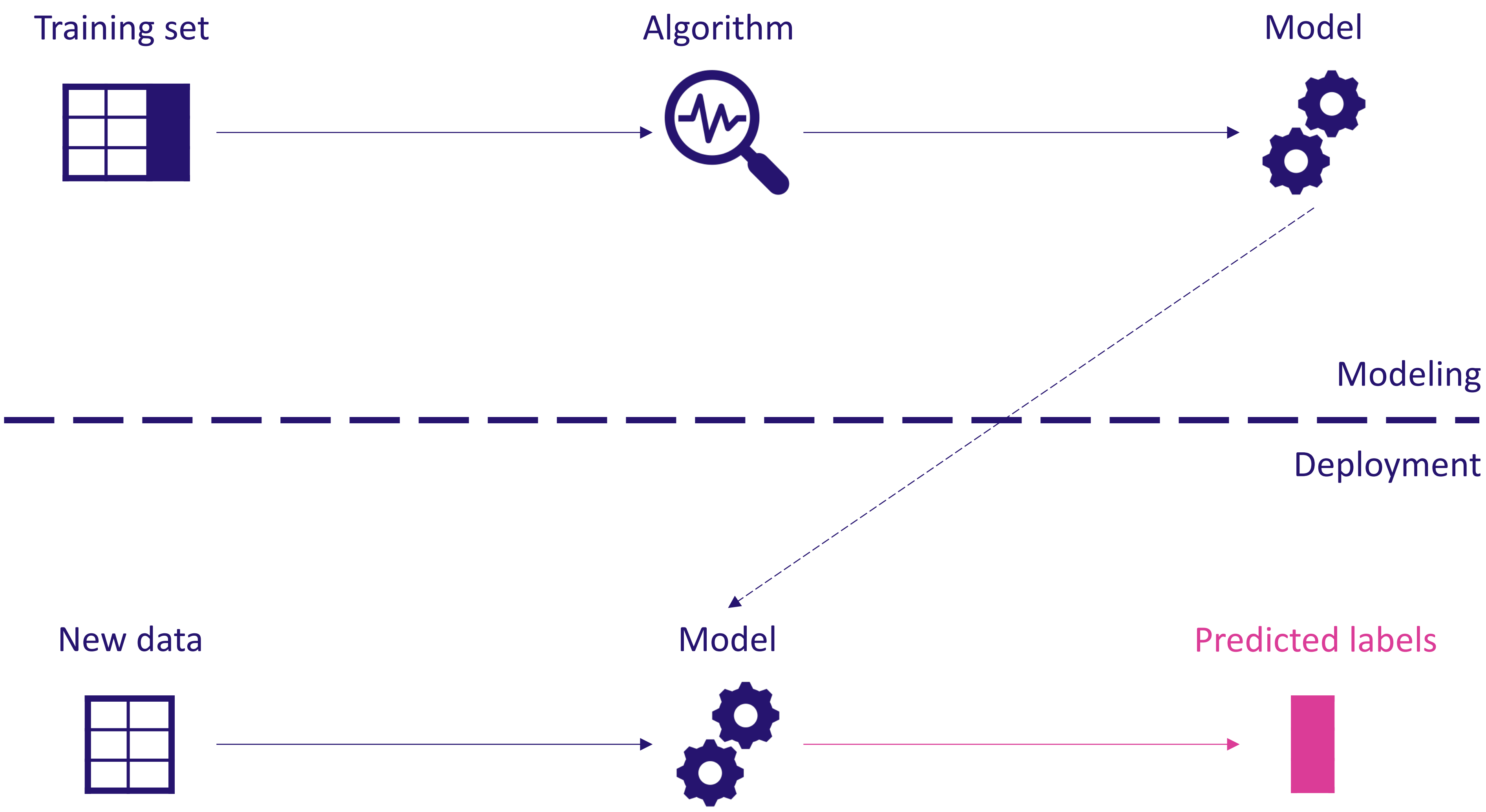
Deployment



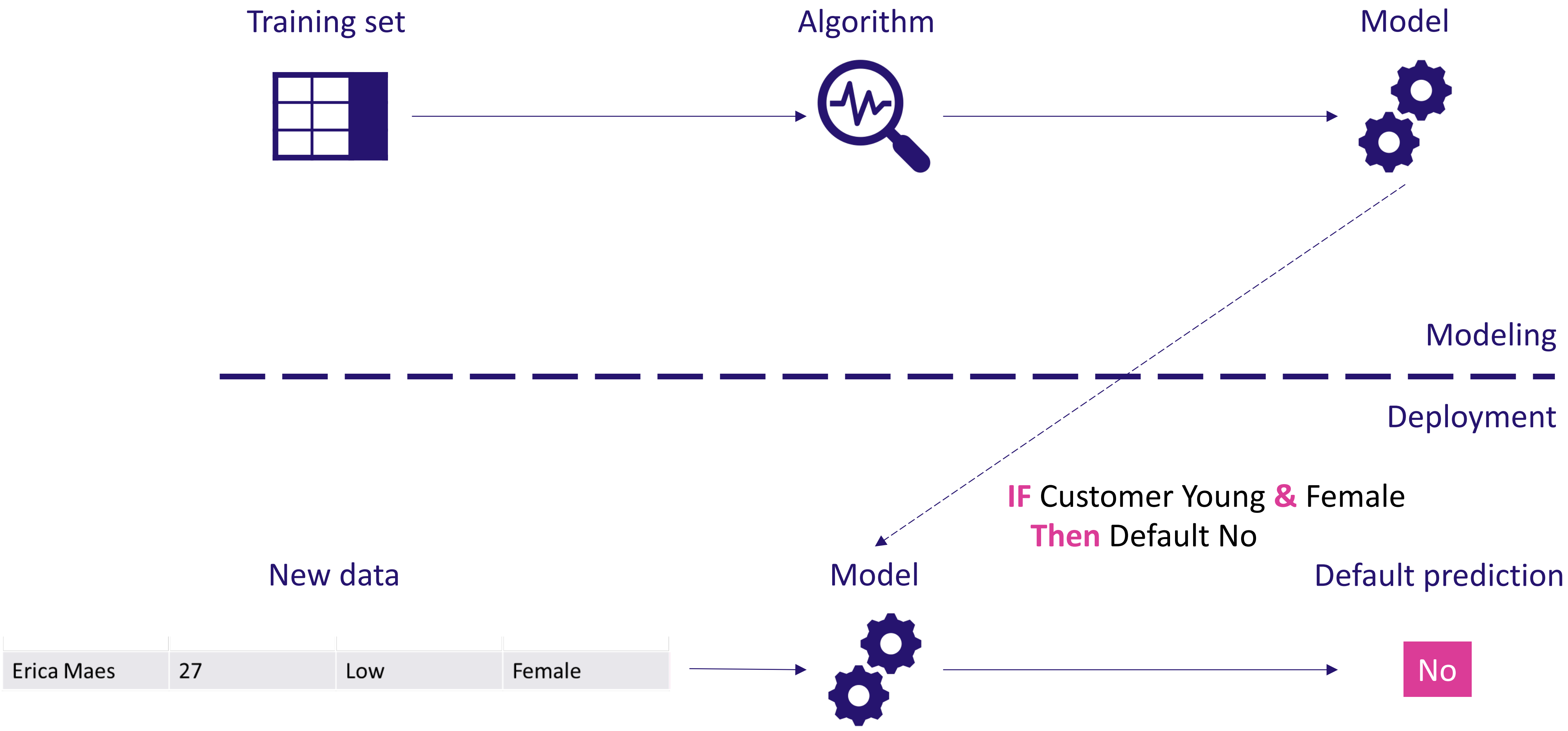
Deployment


Training set
Algorithm
Model
Modeling
Deployment
New data
Model

Deployment


Training set
Algorithm
Model
Modeling
Deployment
New data
Model
Predicted labels

Deployment


Training set
Algorithm
Model
Modeling
Deployment
IF Customer Young & Female
Then Default No
New data
Model
Default prediction
Erica Maes
27
Low
Female
No

4 Evaluating a project

AI business impact

Return on investment (ROI)

- Compare tangible value and cost
- KPI examples
  - Revenue increase
  - Efficiency gains
  - Implementation costs
  - Risk (investment, acceptation, regulation, cyber, etc.)

Value on investment (VOI)

- Also consider intangible value
- KPI examples
  - Customer satisfaction
  - Employee motivation
  - Brand reputation
  - Intellectual property
  - Partnerships

How to measure success?

- Project fails if it doesn't meet predefined objectives
    - Define goals upfront at start
    - Well defined, quantifiable and measurable
    - Compare against status-quo benchmark

- Problem definition as a prediction task
    - NOT "stop customers from leaving"
    - BUT "predict which customers will churn and primary reasons of churn"
    - Makes it easier to find technical metrics

Technical vs business metrics

- Always align business objectives with technical metrics
  - Not doing so will yield disappointment
- Define quantifiable business value
  - Translate it into proper technical metrics
- Choose something meaningful to measure in business terms
  - Make their link to the outputs of the AI solution explicit

Case studies

- Fully automated image clean-up for a manufacturing company
  - Background removal
  - Time and resource consuming

Time to clean an image **0.07s**

Cleaning pixel accuracy **> 99%**

Man hours saved per year **> 3500**

- Physical advertisement impact prediction
  - Assess the effectiveness of ads

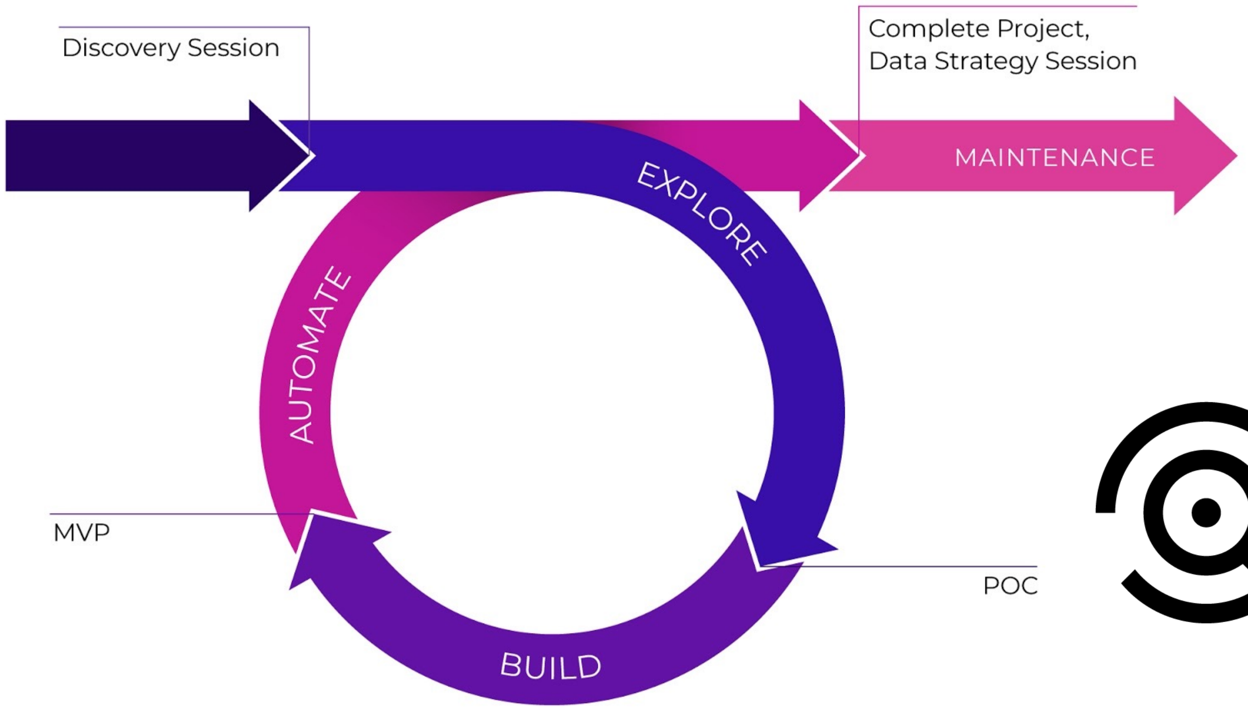
3 billion data points processed in under **10 min**

prediction performance **f1 ~0.8**

Algorithm used in **all reports**

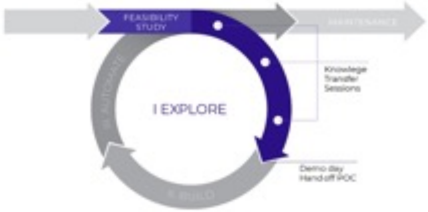
5 Practical AI solutions

Explore - Build - Automate



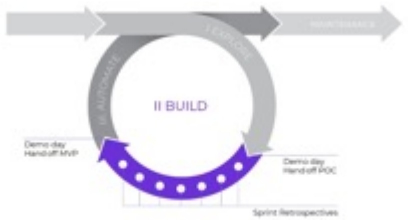
PROPHECY LABS

Explore



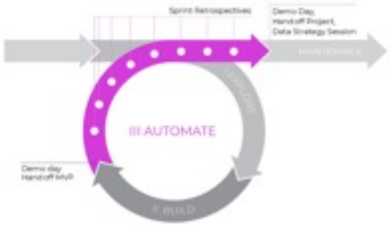
- Test use case viability
  - Worth investing time and money into technological solution?
- Provide minimal solution with a Proof of Concept (PoC)
  - Compare different approaches
  - Identify potential problems + efficient solutions
  - Test software, hardware and infrastructure
  - Gain essential experience, skills and confidence
- Convince stakeholders to invest in AI solution

Build



- Develop and deploy a fully functional Minimal Viable Product (MVP)
- Provide tangible business value
  - Actionable insights, forecasts, optimized processes or automated tasks
- Satisfy all requirements with focus on
  - Stability, testing and good code quality
  - Proper infrastructure
  - Scalability if needed

Automate

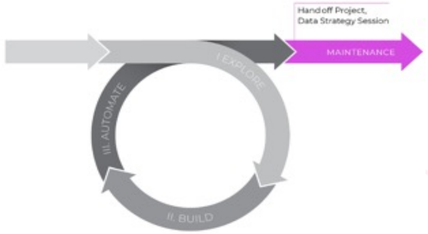


- Make the software operate autonomously
  - Minimal human supervision needed

- Integrate the solution in current infrastructure
  - Fully scalable and robust

PROPHECY LABS

Maintenance



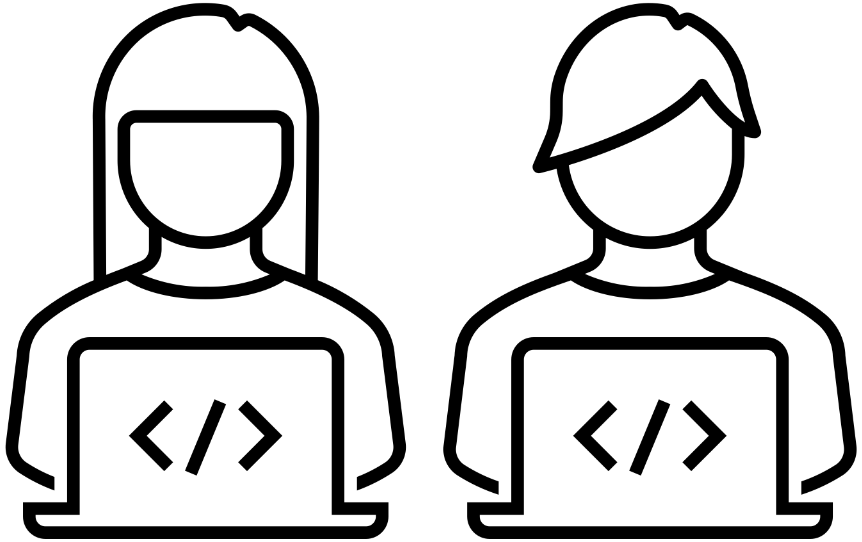
- Evaluate and monitor performance of the model
- Track model quality and predictions over time
- React to feedback from deployed models

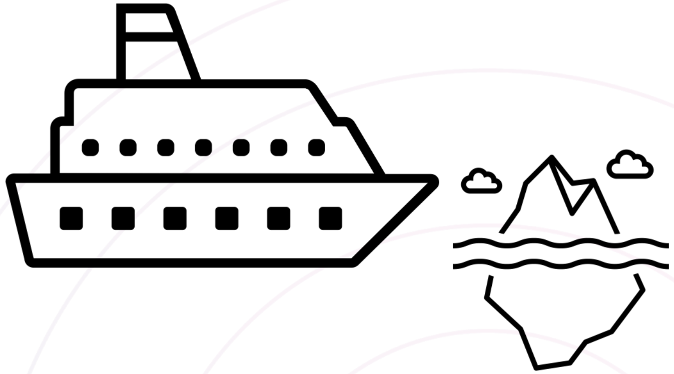
nannyML

6 Hands-on tutorial

Jupyter Notebook

Happy coding!





PROPHECY LABS

ctg

nannyML

AI4Business

ESF

INVESTEERT IN JOUW TOEKOMST

Europese Unie

Vlaanderen is werk