




zalando

Predicting Customer Lifetime Value

Alice Lynch

Research Engineer

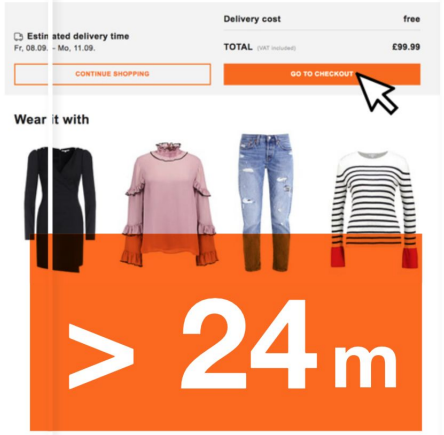
Zalando

Datafest Tbilisi 2018

- 2,300 orders / minute
- 1.5 million shipments
- 1,200 shipments / minute
- 72,000 shipments / hour

zalando

Active Customers



> 24 m

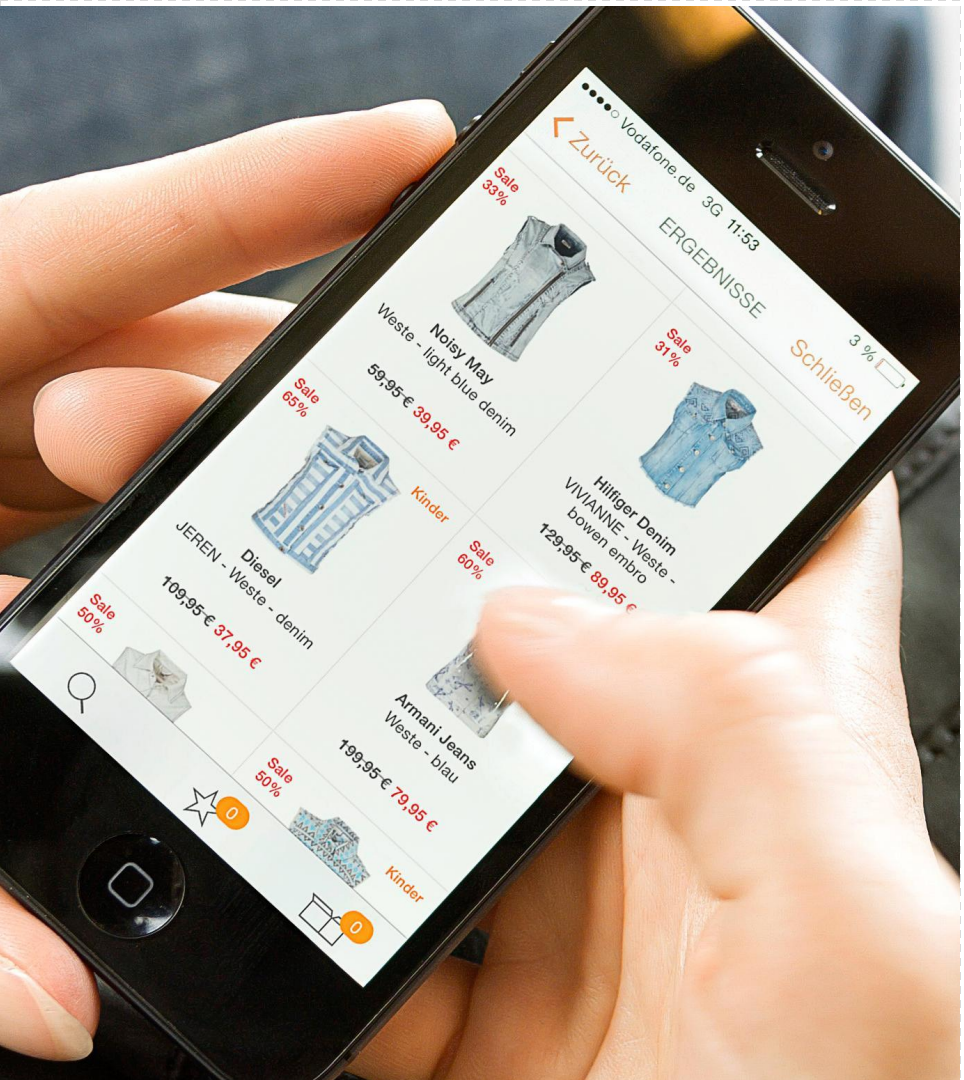
Assortment

> 300,000

Markets

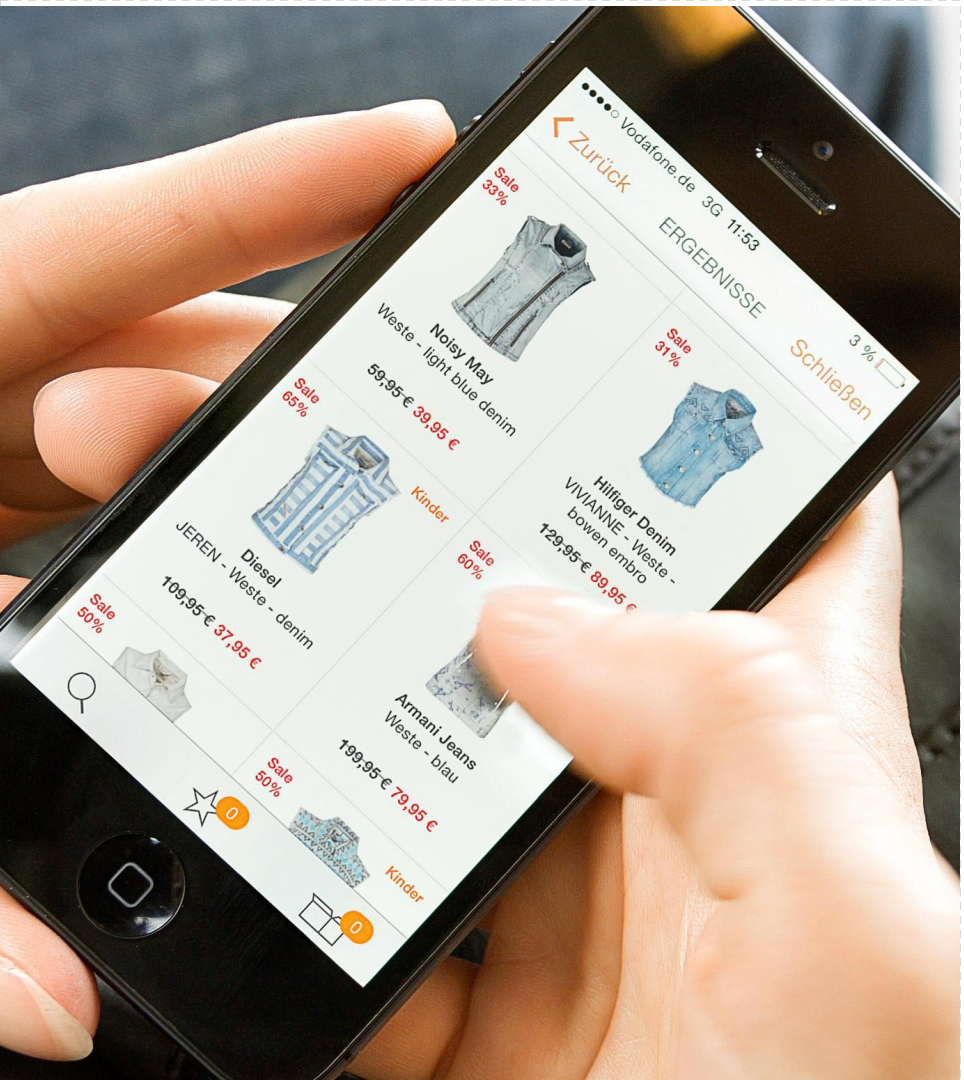
~ 2,000

Brands



What is Customer Lifetime Value (CLV) ?

- **Total profit of entire relationship with customer**
- Focus on the long term health of the customer relationship
- Cost to attract, serve & maintain
- Value of transactions

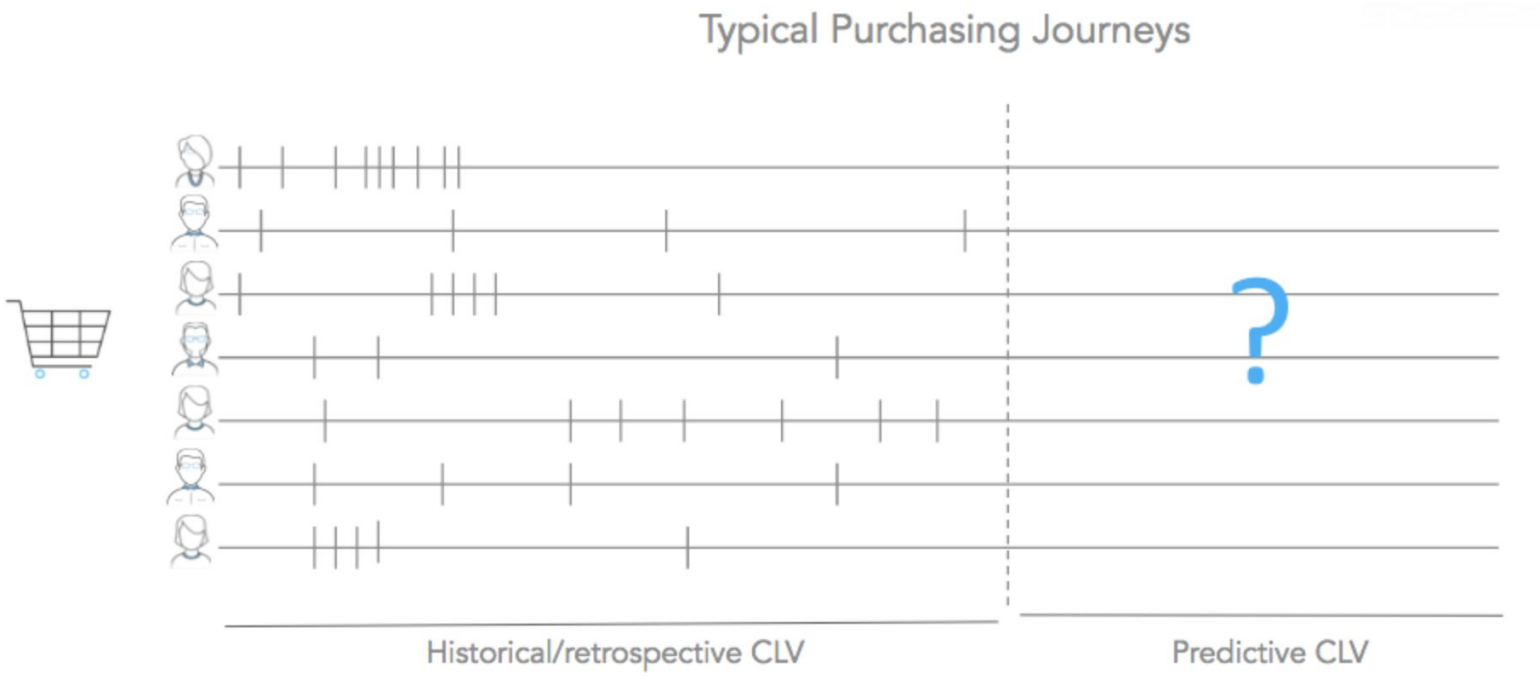


Why use CLV?

- Customer Segmentation
- Identify traits of Valuable Customers
- Allocate limited resources (logistics, support etc.)
- Costs to acquire & maintain

CLV: A HISTORICAL HEURISTIC APPROACH

- Sum the profit to date generated by each customer as assign this as CLV



CLV: COMPARING APPROACHES

	Contractual	Non-Contractual
Continuous Purchases	• Company knows when customer leaves • Customer can buy any time • *Credit Cards* Probabilistic Models: Exponential-Gamma	• Company does not know when customer leaves • Customer can buy any time • *Online Clothes Retail* Probabilistic Models: Pareto/NBD R package 'Buy Til You Die' **Machine Learning Models** Asos Paper Groupon Paper
Discrete Purchases	• Company knows when customer leaves • Customer buys at regular intervals • *Gym membership* Hierarchical Bayes Model	• Company does not know when customer leaves • Customer buys at regular intervals • *Buying Prescription* Probabilistic Models: BG/BB Model

CLV: A MACHINE LEARNING APPROACH

TRAIN A MACHINE LEARNING MODEL:

- Use historical data to learn the relationship between past customer behaviour and future customer value

SCORE PRESENT DATA WITH TRAINED MODEL:

- Predict customer value for future time frame

+ **DEVELOP RICH PREDICTIVE FEATURES**

+ **CHOOSE DIFFERENT ALGORITHM APPROPRIATE FOR YOUR DATA DISTRIBUTION**

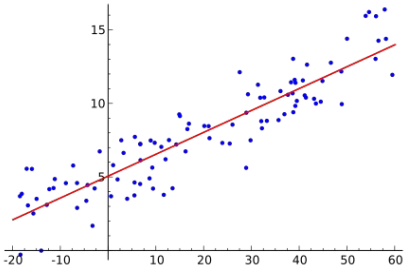
- **TRAINING DATA NEEDS TO BE OVER LONG TIME FRAME**

- **CANNOT PREDICT INFINITELY INTO THE FUTURE**

CLV: KNOW YOUR PROBLEM

TRANSLATE YOUR BUSINESS USE-CASE INTO A MACHINE LEARNING PROBLEM

BUSINESS CASE		MACHINE LEARNING APPROACH
RANKING CUSTOMERS BY CLV	→	LEARN TO RANK
SEGMENTING CUSTOMERS BY CLV	→	MULTI-CLASS CLASSIFICATION
PREDICTING ABSOLUTE VALUE OF CLV	→	REGRESSION



zalando

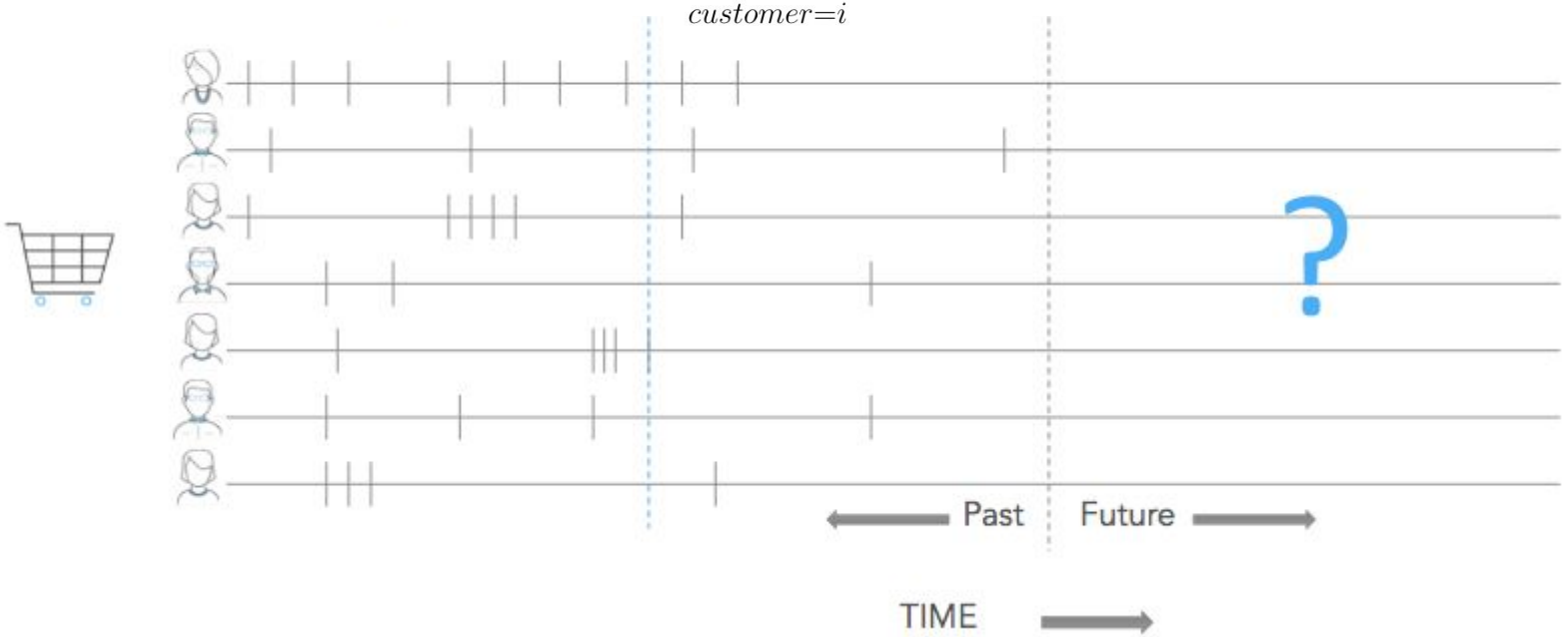
CLV: KNOW YOUR PROBLEM

FEATURES

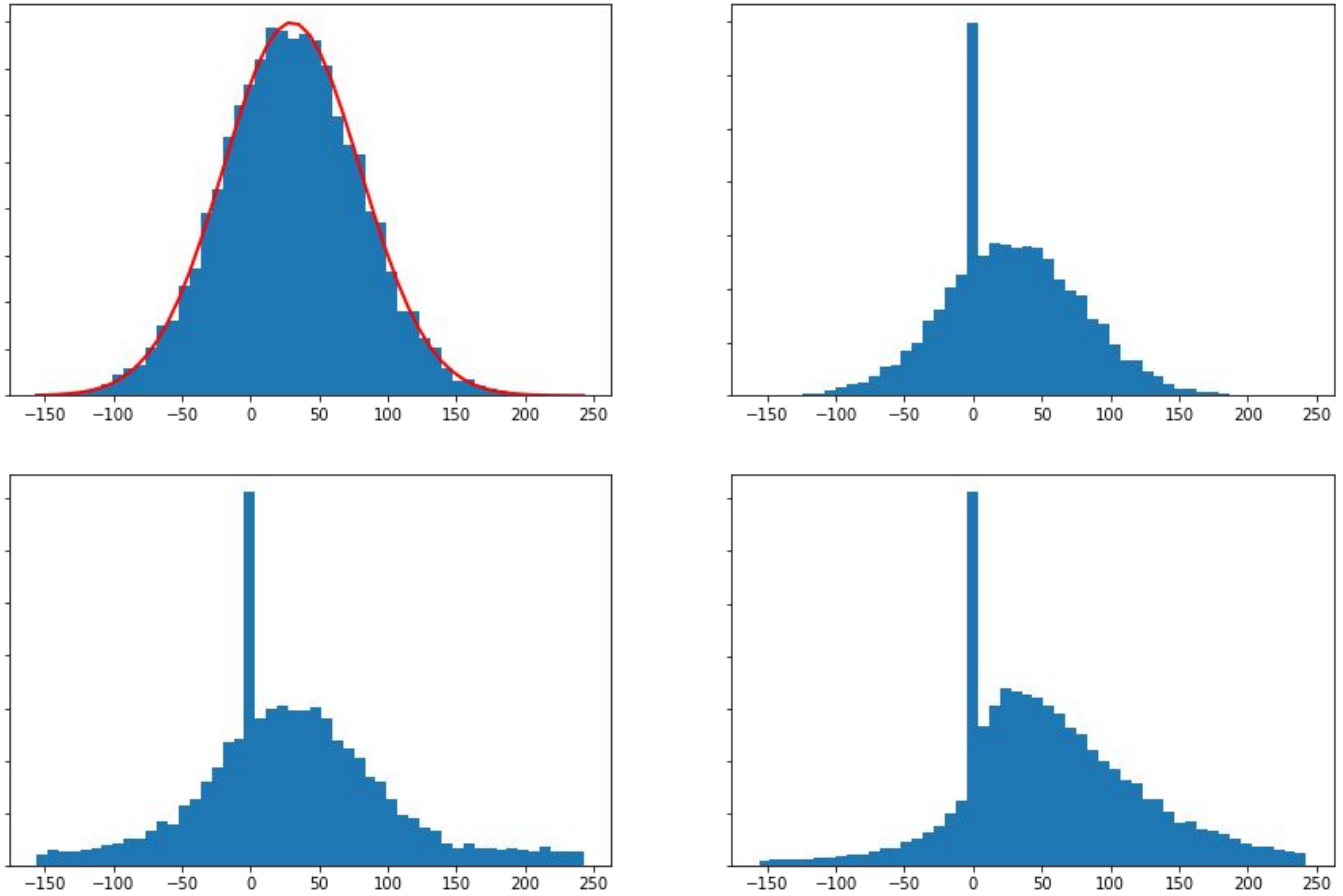
TARGET VARIABLE

$$\sum_{customer=i}^{n} profit_i$$

FORECAST PERIOD

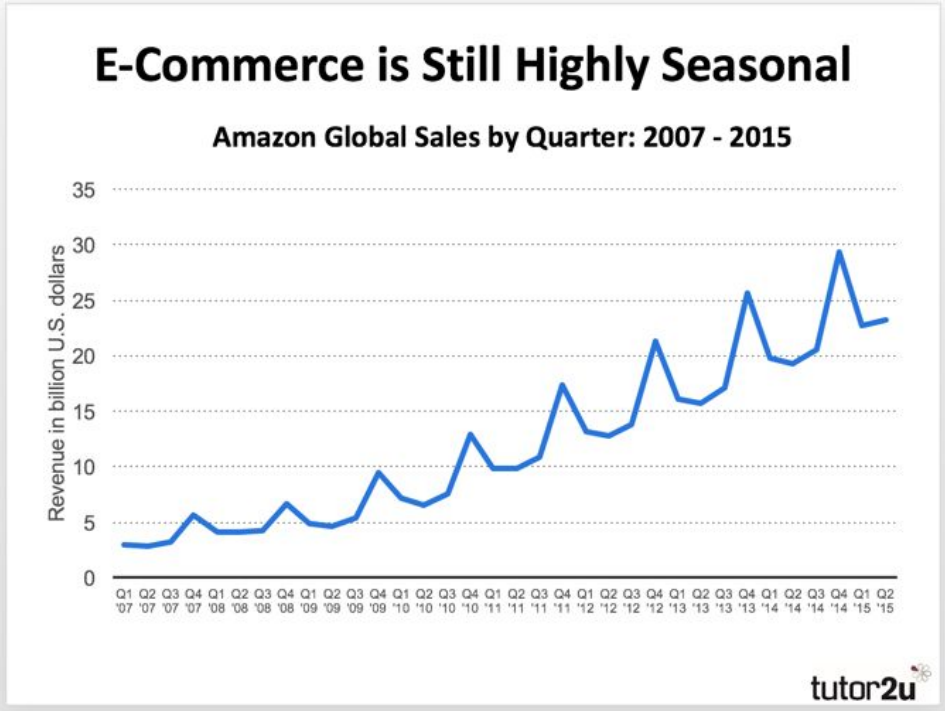


CLV: KNOW YOUR DATA



zalando

CLV: KNOW YOUR DATA



STRATEGIES FOR DEALING WITH SEASONALITY

- Seasonal Features
- Time of year/month features
- Model a non-season time frame, e.g. 1 year
- Re-train frequently

zalando

FEATURE ENGINEERING

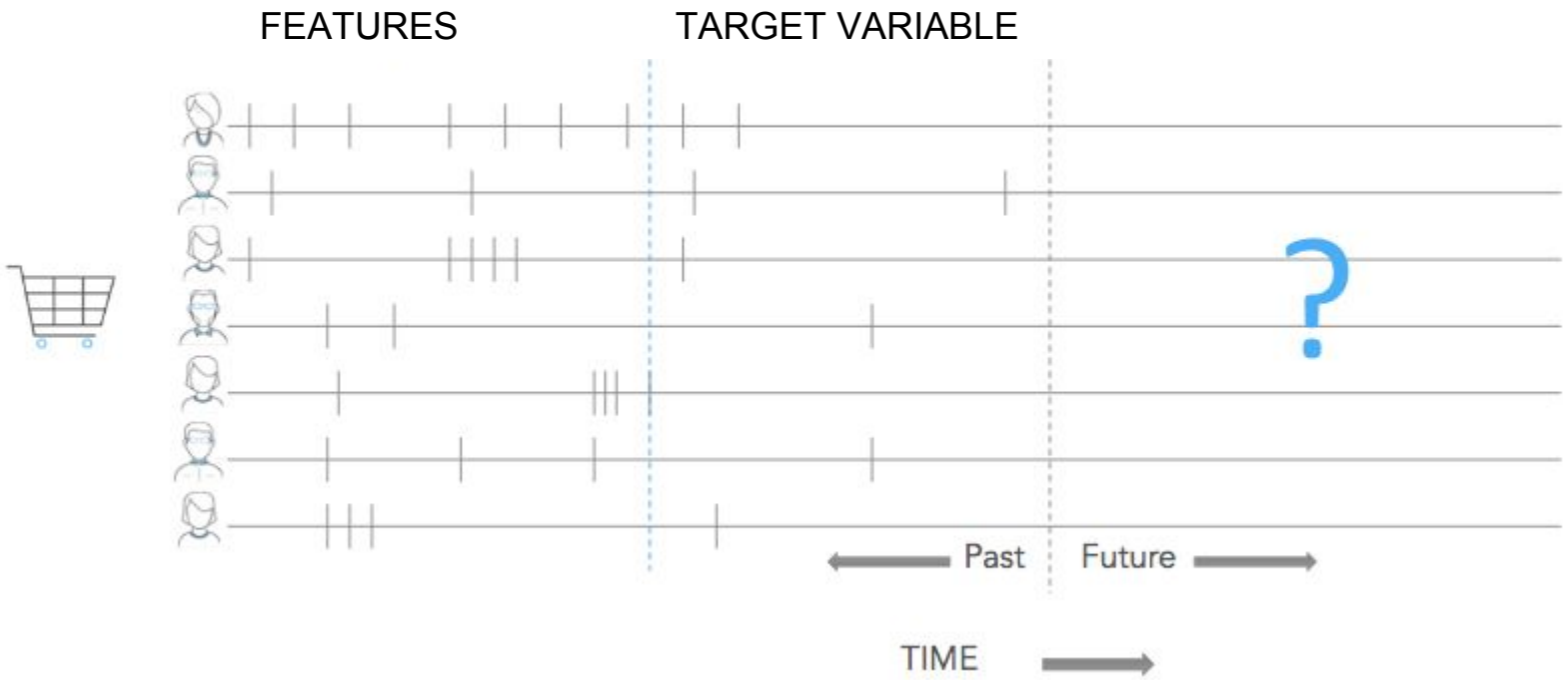
“Features” are the Foundation of any Machine Learning Model

Capture customer behaviour that you believe will predict customer value

- Use insights from your business knowledge
- Recency, Frequency, Monetary Value
- Include Seasonality if it affects your business

FEATURE ENGINEERING

“Features” are the Foundation of any Machine Learning Model



FEATURE ENGINEERING

“Features” are the Foundation of any Machine Learning Model

Id	Features				Label
customer_id	Number of Purchases Last Year	Profit from two quarters ago	Fraction of items on sale	...	Profit in 2018
123	2	32.76	0	...	74.12
456	0	-12.23	0.87	...	-20.10
...					

FEATURE ENGINEERING

“Features” are the Foundation of any Machine Learning Model

Reliability: Garbage in, Garbage Out

- What data is available to you reliably?
- Can you access this data in production?
- Are there issues with Data Quality?

Know your Features:

- Is there a relationship with the label? (correlation, chi squared)
- Are they highly correlated with each other?

MODELLING

Train an algorithm that estimates the relationship between features & label

$$CustomerValue = f(CustomerBehaviour)$$

LINEAR REGRESSION

+ FAST TO TRAIN
+ EASY TO INTERPRET

- UNDERLYING DATA ASSUMPTIONS
- DATA TRANSFORMATIONS

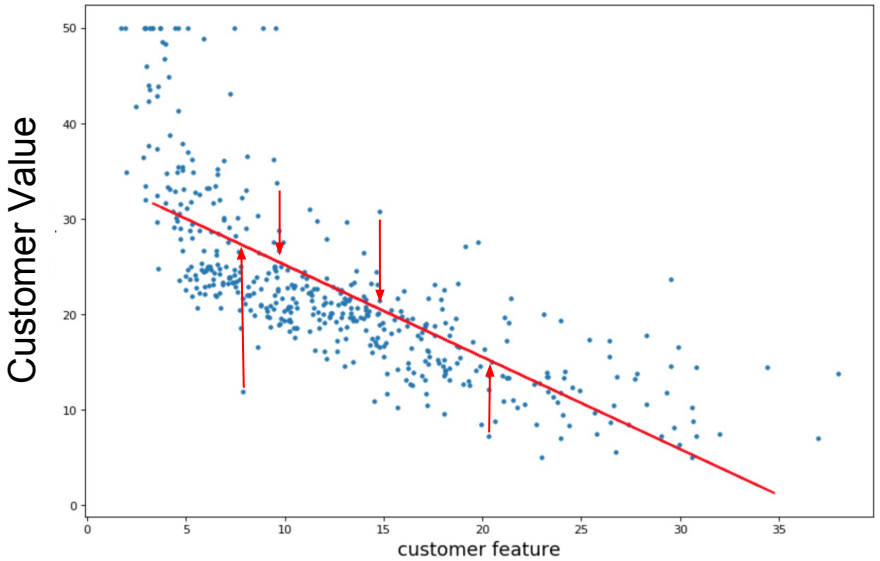
NON-LINEAR ALGORITHMS

+ LOTS OF METHODS FOR DATA TYPES
+ LESS DATA ASSUMPTIONS

- SLOW TO TRAIN
- HYPER-PARAMETERS TO TUNE
- HARDER TO INTERPRET
- MAYBE HARDER TO DEPLOY

LINEAR REGRESSION

$$\text{label}_i = f(features) = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} +$$



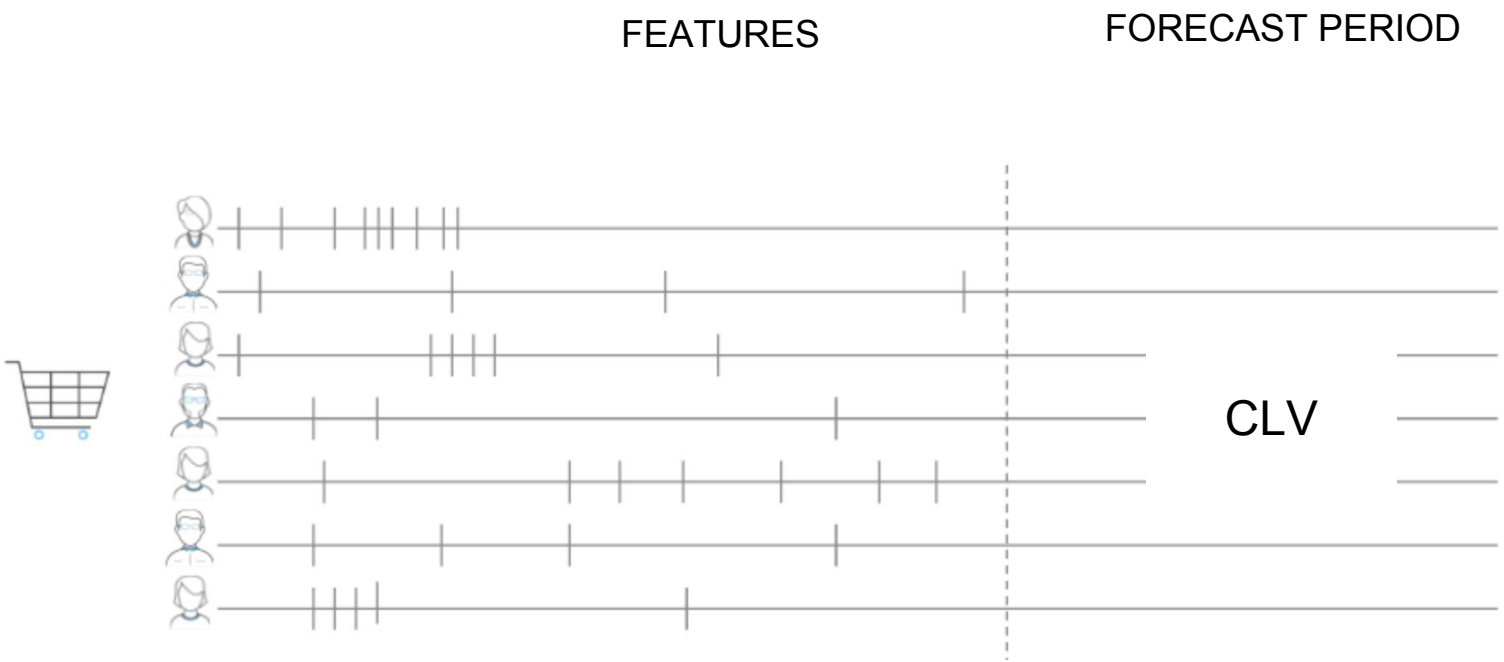
Minimize the Mean Squared Error:

$$\text{MSE} = \frac{1}{n} \sum_{i}^{n} (f(x_i) - y_i)^2$$

MODEL EVALUATION

- **Use a Hold-out Test Set** → *Know if you are Overfitting*
- **Technical Performance:**
  - MSE needs a baseline → *How good is good enough?*
  - Calculate MSE for your baseline / heuristic
- **Business Performance:**
  - Track impact on KPIs → *Is there good Model-Business Fit ?*
  - Profit, revenue, NPS, Shipping time for VIPs

PREDICTING



zalando

QUESTIONS???

DEMO