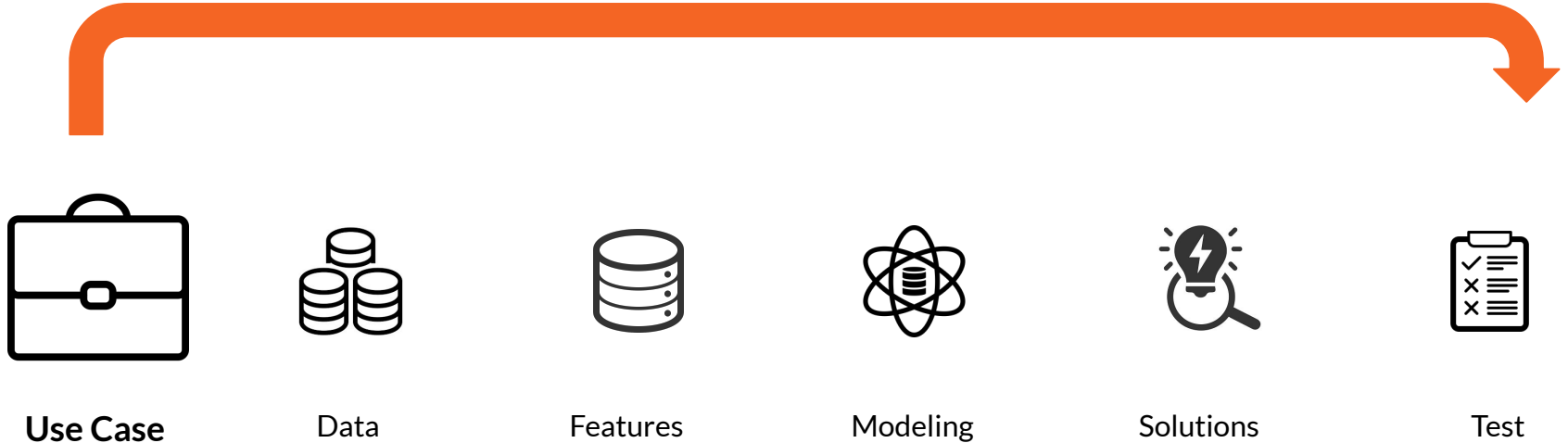

Churn Prevention - Telecom Industry

Machine Learning Proposal

Churn Prevention Process



2 weeks

Business Problem



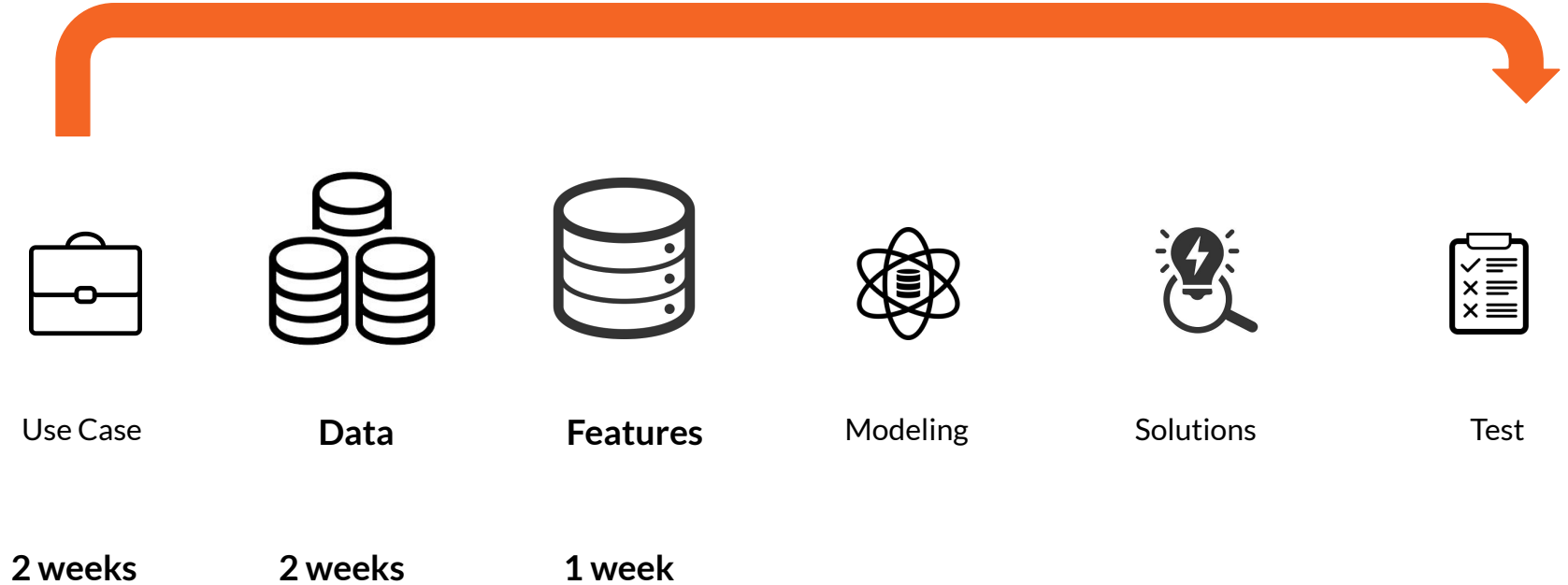
The churn of initial three months is high for newly acquired customers



Data Science Use Case

1. Gather data & builds a model - a classifier that predicts churn
2. Devise solutions to re-engage the nearly-churning customers
3. Test the effectiveness of the solution

Churn Prevention Process



Data & Features



The machine learning model is only as good as its data source

- Data collection & cleaning
 - Understand the **context** - customers, plans, competitions
 - Gather **relevant** data, as much as possible
 - Make sure the data is **complete, clean** and **accurate**
- Feature selection & engineering
 - Include as many **relevant features** as possible
 - Prepare and cleansed data (**scaling, encoding** etc.) for modeling

Data & Features



The machine learning model is only as good as its data source

- Input data
 - **Demographics:** gender, age, address, postcode, marital status, dependants etc
 - **Acquisition:** organic, PPC, display ads, prints, referral, trial etc
 - **Service:**
 - Activate date, plan, tenure, add-ons, pricing, promo, auto-renewal, dock
 - # of calls, call length, # of texts, # of voice message, data
 - payment method, billing, top-up, pay delay, suspension
 - # of calls/emails to customer service, # of complaints, resolved issue, web logs
 - **Technicality:** service coverage, 5G, number of receivers etc
- Output data: **churn** or not

Data & Features



The machine learning model is only as good as its data source

that CAN be modeled:

1. complaint
2. customer service
3. coverage
4. trial period
5. prepaid accounts
6. plan variations

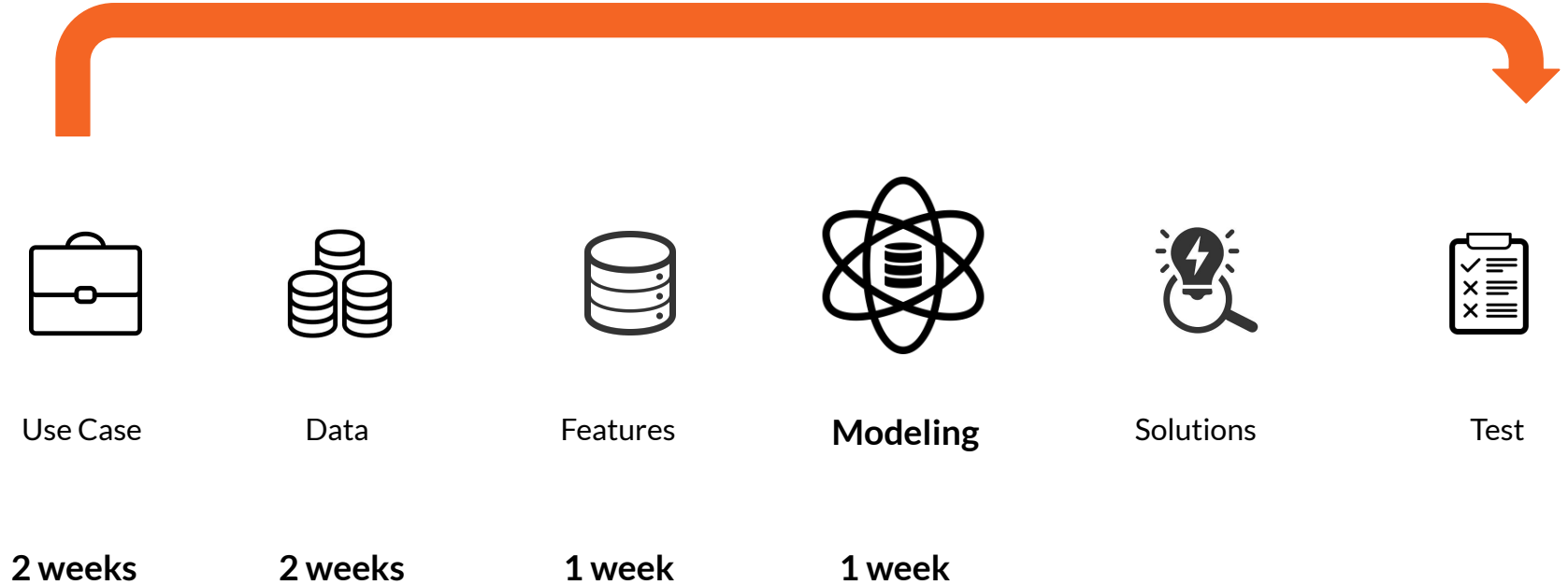
.....

that CANNOT be modeled:

1. silent churn: no sign at all
2. better offers from competition
3. receive a referral to join competition
4. switched to a phone plan
5. naturally churn: use for travel etc.
6. relocation

.....

Churn Prevention Process



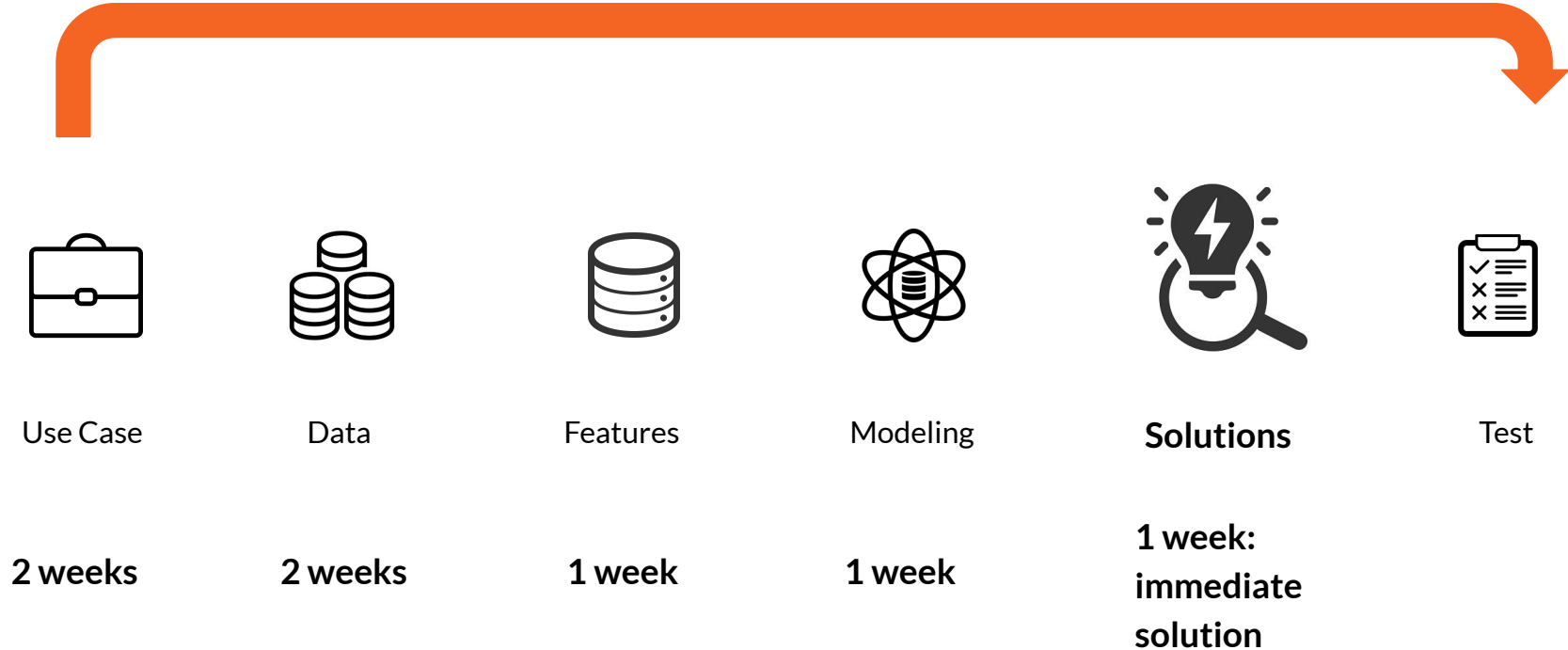
Modeling



Build a model - a classifier to predict churning customers

- Model evaluation
 - Recall: % of churns correctly classified out of the total **actual churns**
 - **How good is the model to identify churning customers**
 - Precision: % of churns correctly classified out of the total **predicted churns**
 - **How accurate is the model to identify churning customers to re-engage**
 - F1 score: harmonic mean of precision and recall
 - **Ideal in this case** as we need to first identify churning customers and then re-engagement
- Model selection
 - Logistic regression
 - LightGBM
 - CatBoost
- Feature Importance: impact of features on churn
 - Identify pain points for churning customers
 - Structural-triggered or accident-triggered?
 - Prioritise to work on the reasons identified

Churn Prevention Process



Solutions



Device a solution balancing immediate and long-term business needs

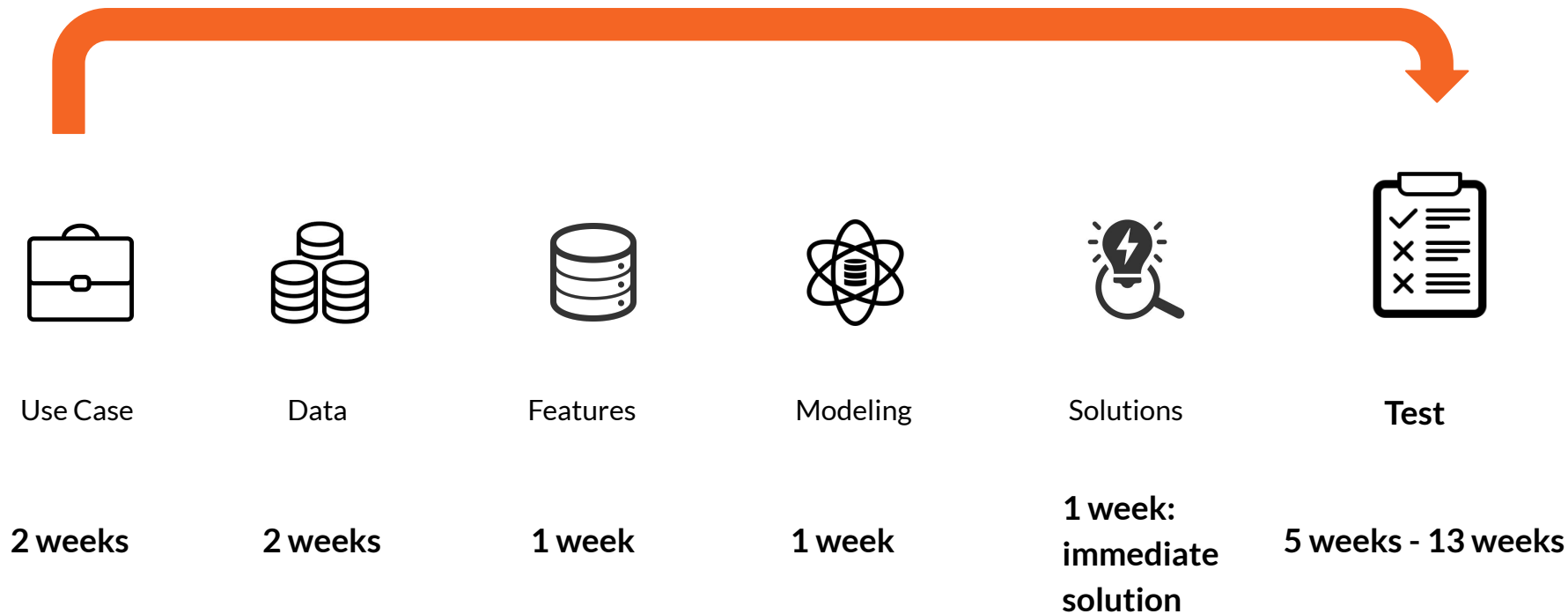
- Offer financial incentives to customers who are about to churn, such as promo code, coupon, bonus etc - **Immediate Solution**
- Calculate customer lifetime value (CLV) of churning customers and offer financial incentives to customers based on their CLV
- Monitor usage and offer them incentives such as call minutes and data, and improved service bundles such as Netflix, Youtube
- Segment customers based on their traits and usage etc, and design flexible plans - **Ultimate Data Science Solution**
- However, **retention strategy should not only focus on *win-backs*** but requires **well-rounded CRM** from acquisition, communication, management to churn prevention

Low-effort: **blanket** solution for all churning customers



High-effort: **customised** solution depending on customer characteristics

Churn Prevention Process



Testing



Design A/B testing to measure effectiveness of the solution

- Experiment design
 - Run the experiment for three months
 - Randomly select new customers into two groups - **70% control** vs **30% experiment**
 - Measure average **CLV** and **churn rate** of two groups at each month
 - Conduct **t-test** to see if the observed difference is **statistically significant** between two groups and gauge the impact of re-engagement solutions by calculating the **confidence interval**

Refine and repeat

Test, measure, and test again...ad infinitum

