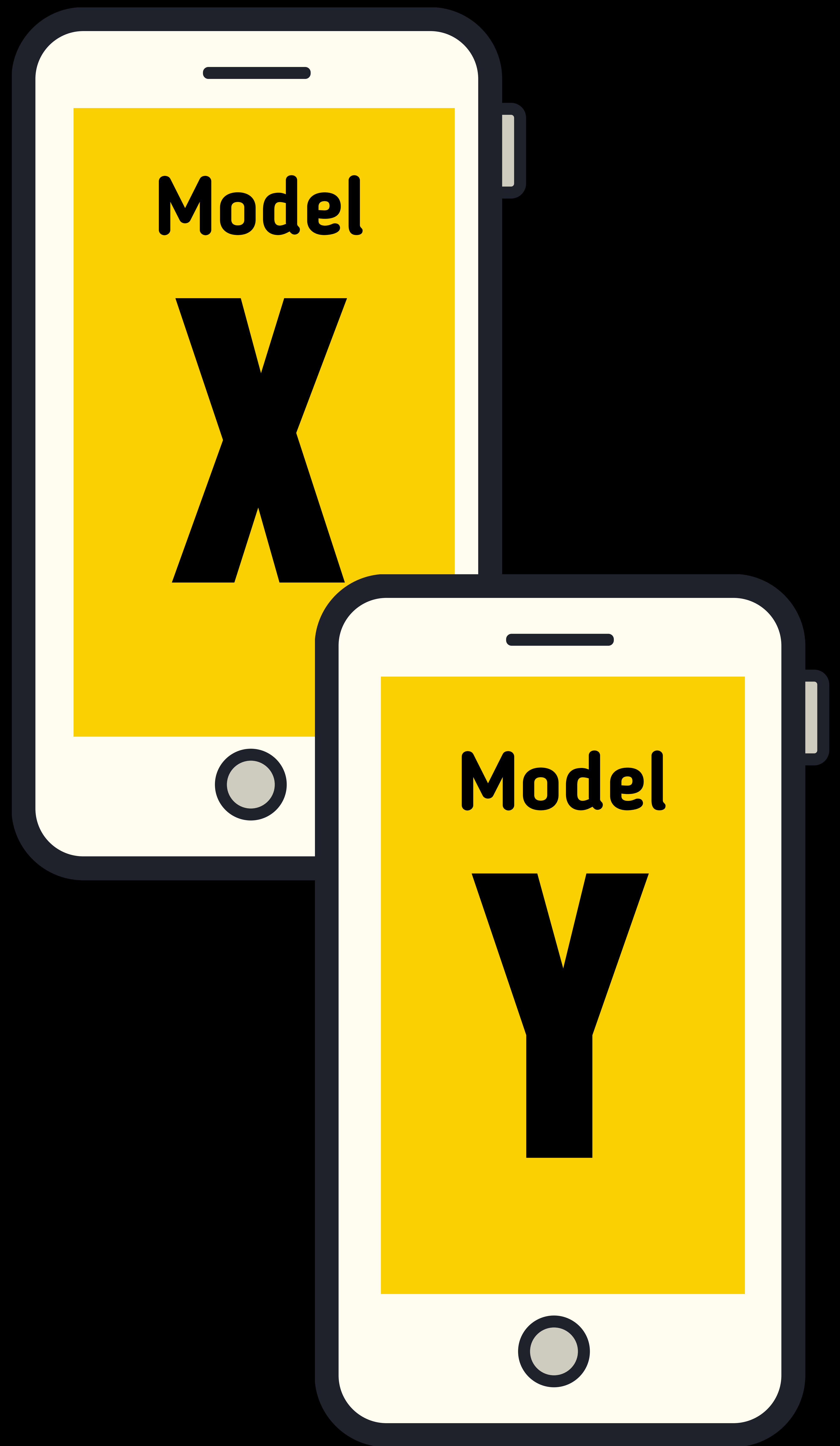
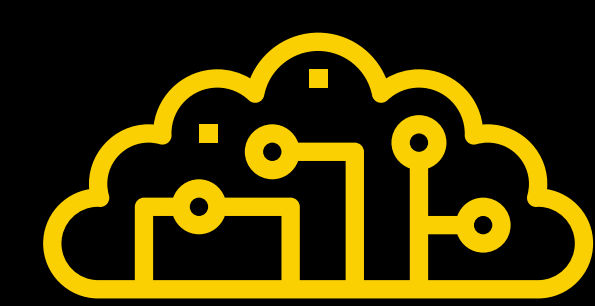


A/B Testing Your Machine Learning Models





Hey manager I developed a new model to improve the recommendation engine

Ok, let's test it on a sample of users

Why sample, why can't we test it on all users so it will be easy to analyse the results

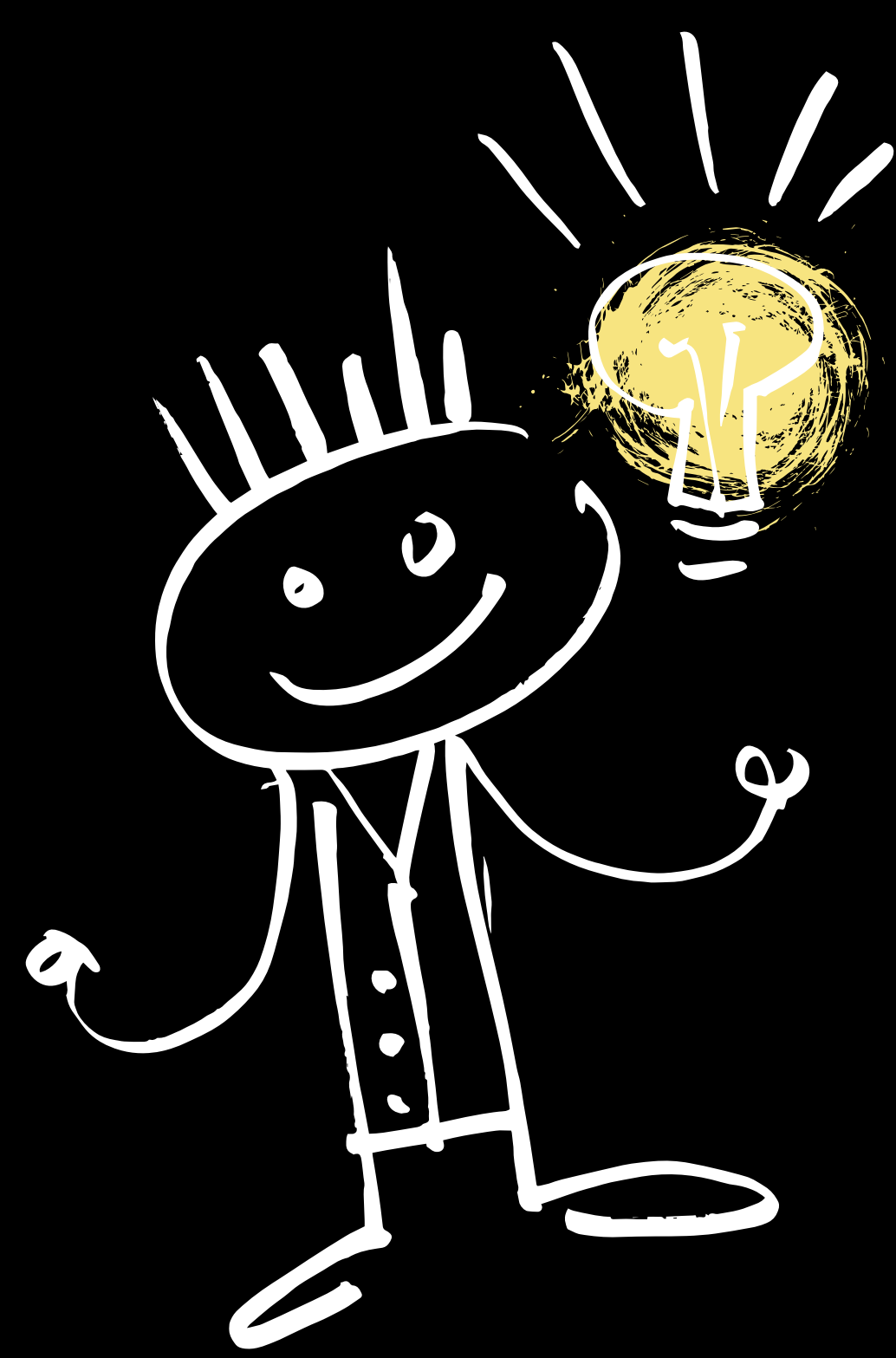
What if it doesn't attract customers and we lose all the trust and revenue as well

So what we should do now to test it

Let's do A/B testing of your model

What is A/B testing, I never heard about it.

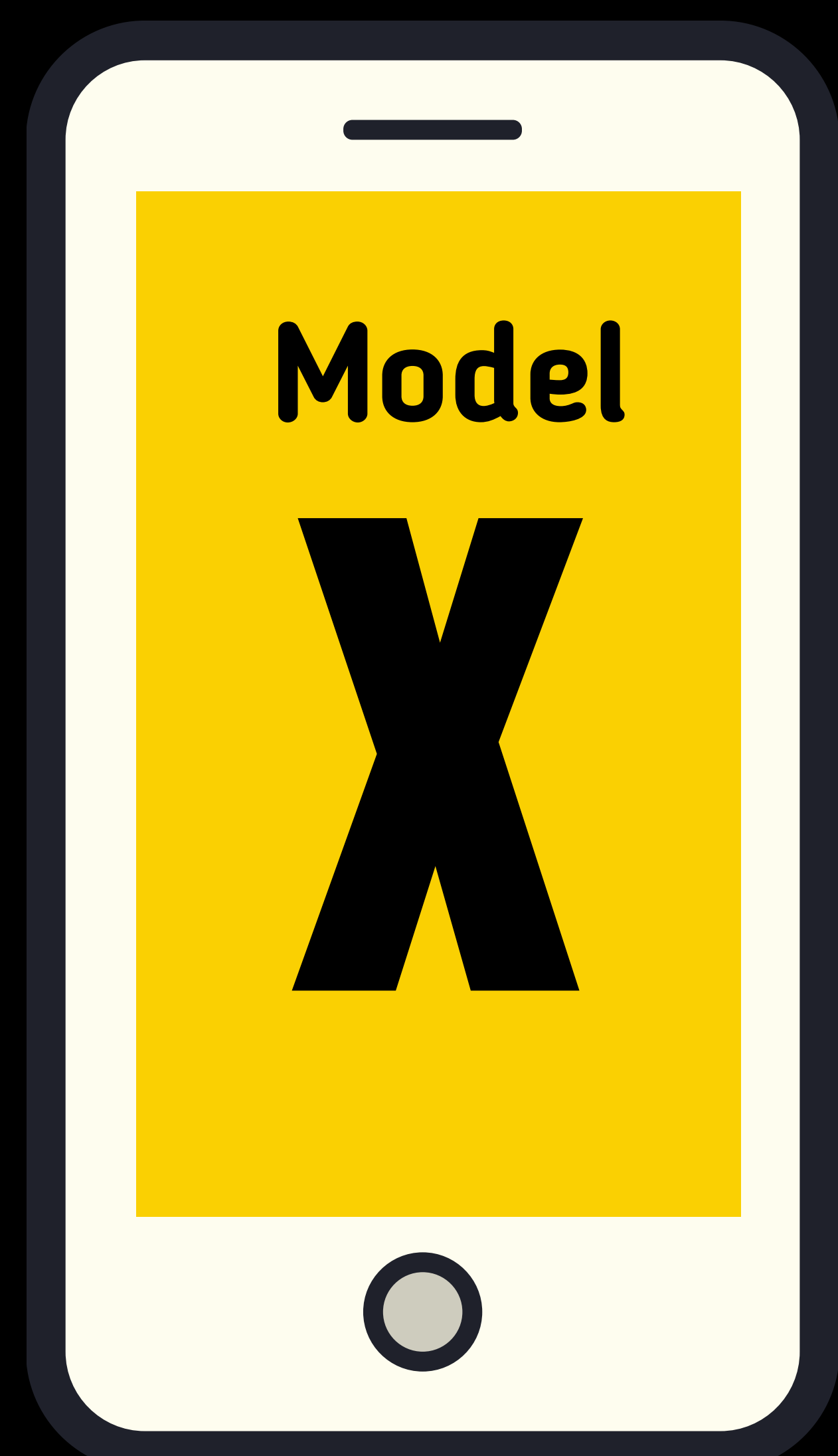
Check the next slides to understand it



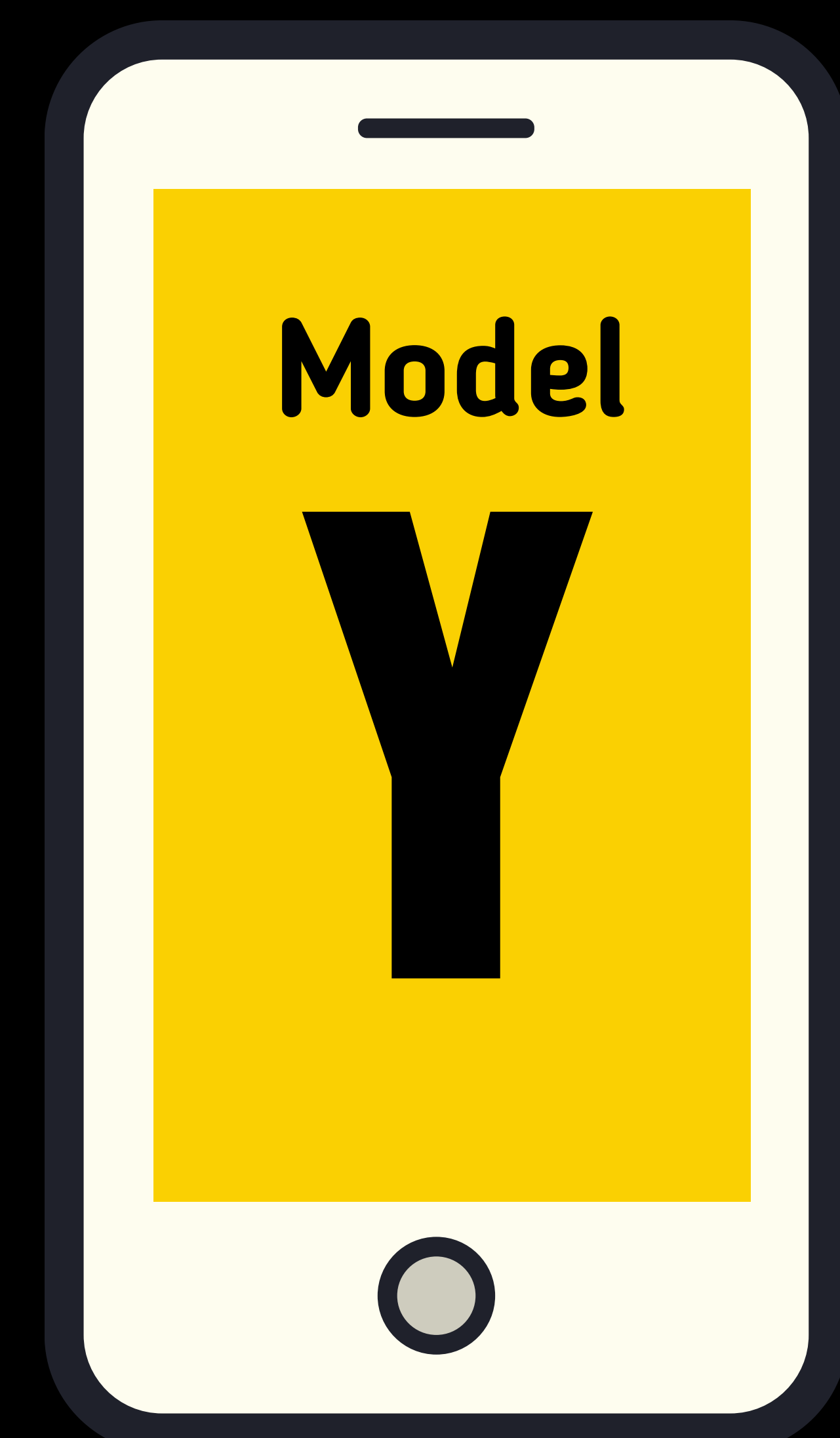
New Employee



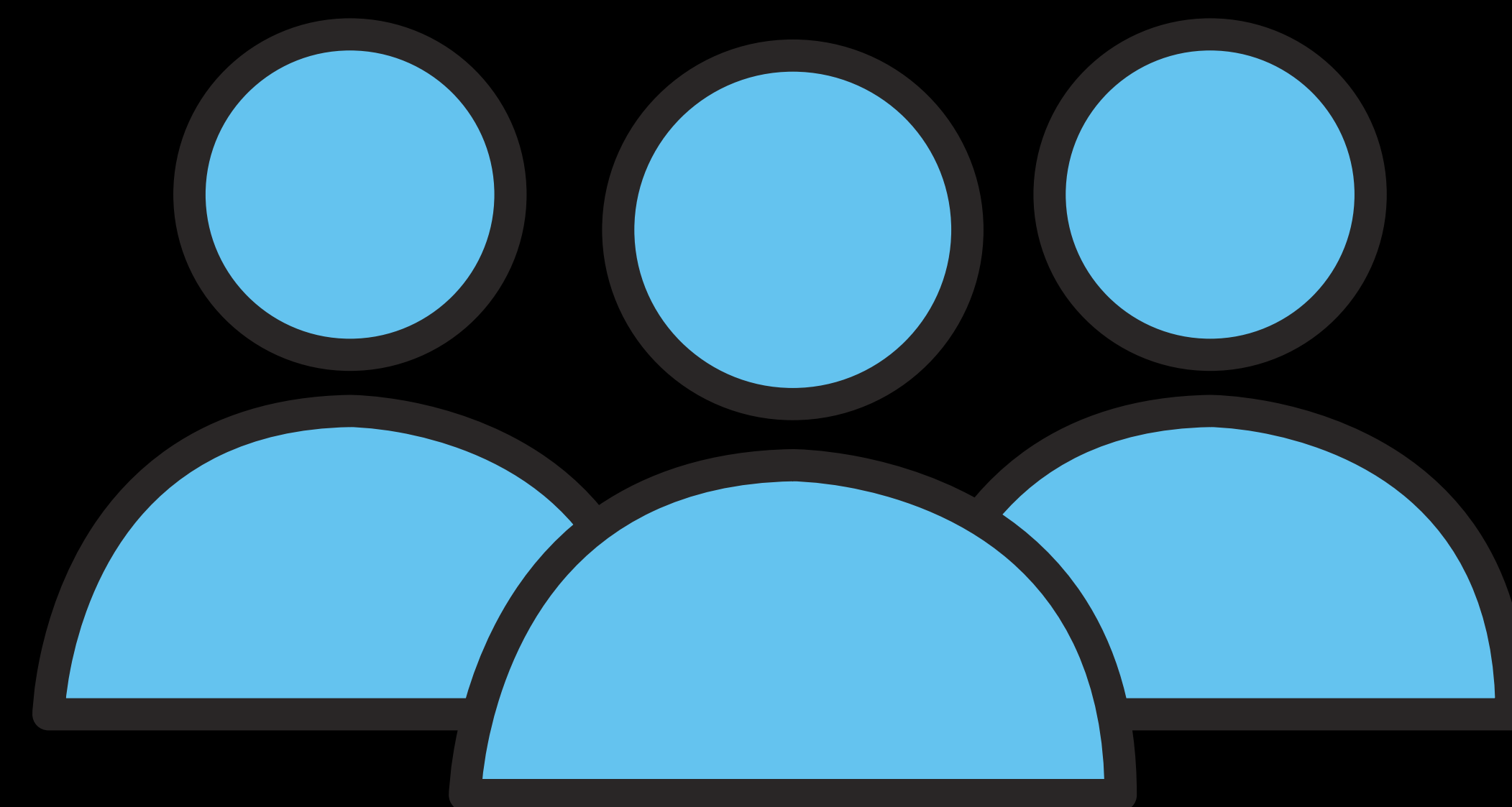
Manager



X = Current model
Y = New Model



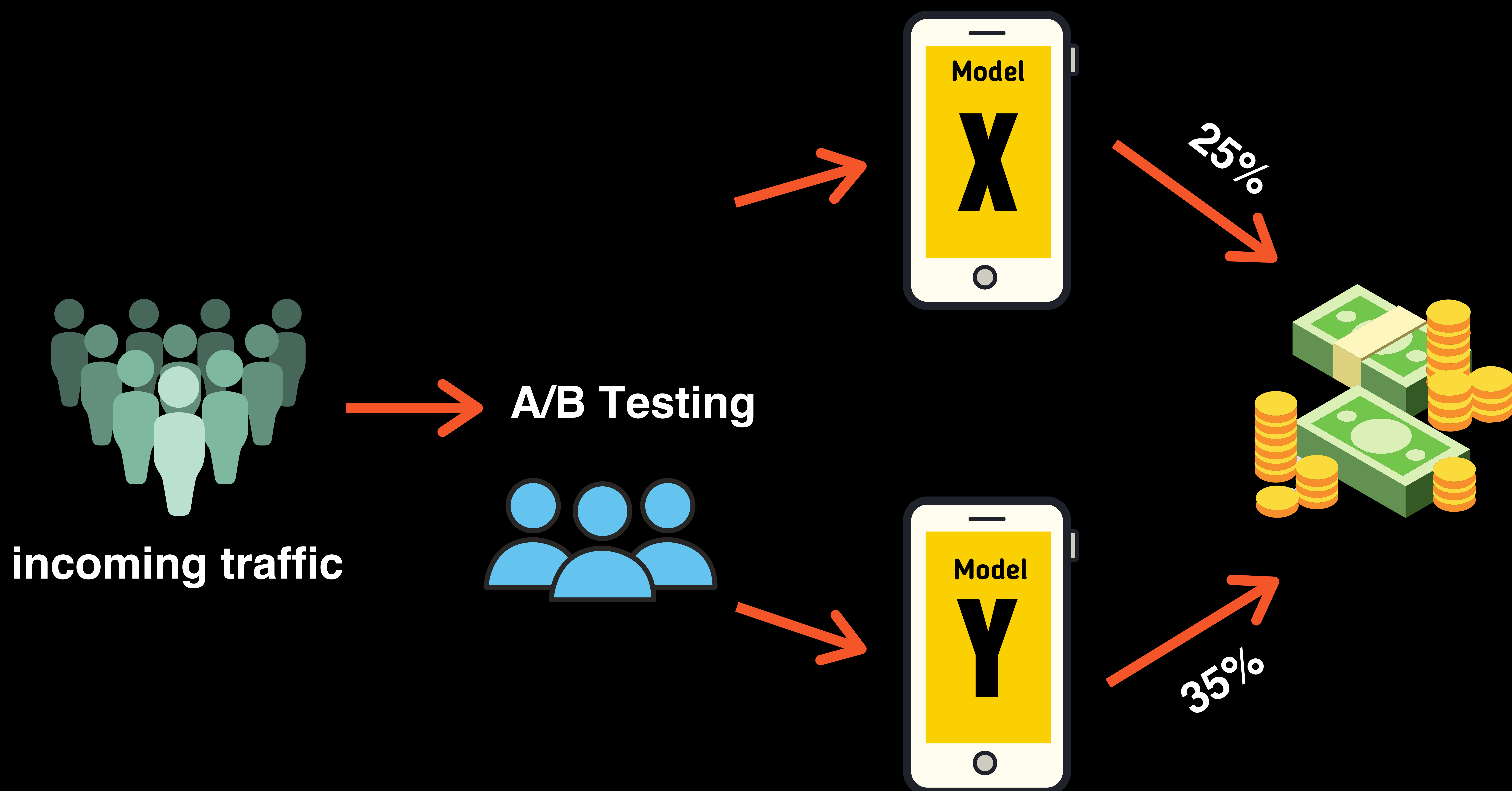
Group - A (Control group)



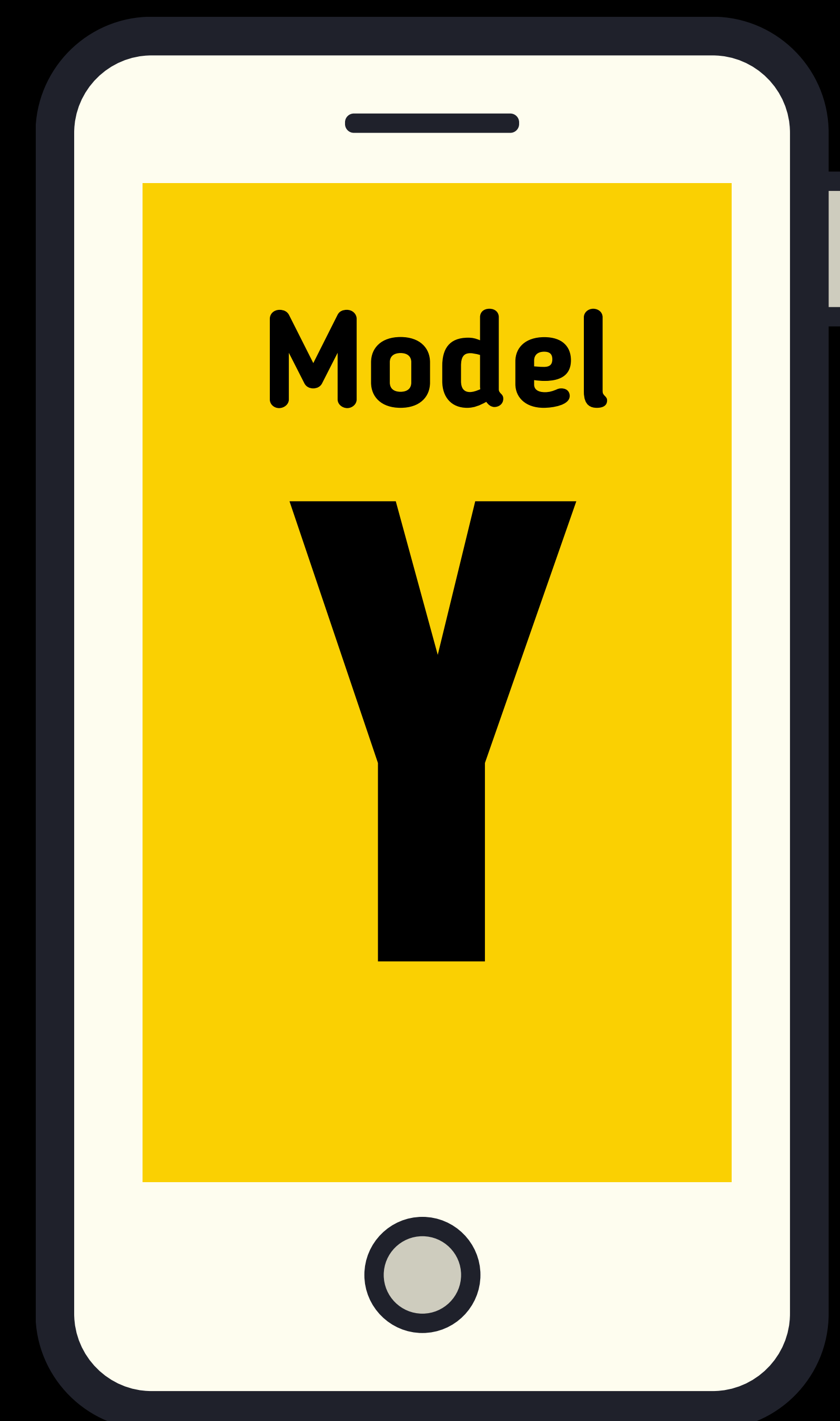
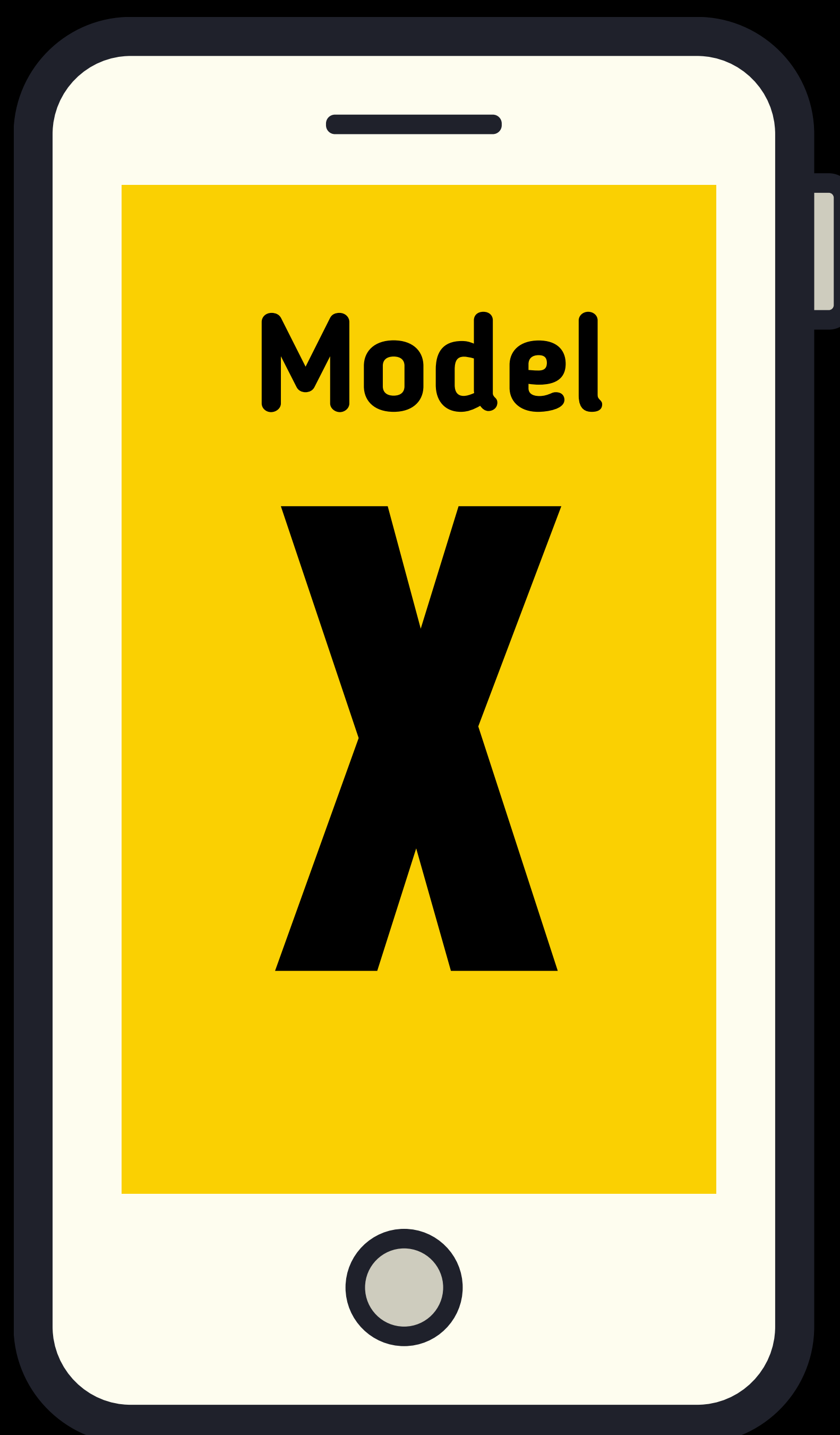
Group - B (Test group)

What is A/B testing

This test enables us to determine whether changes in the model (in our case) improve the click-through rate or purchasing behaviour of a customer (metric in our case) when compared to the current model in use.



Let's say we divided the incoming traffic of users randomly into 2 groups and each group will see different versions of the new model we created and finally we calculate the purchasing behaviour of the user and we can decide which model works best.

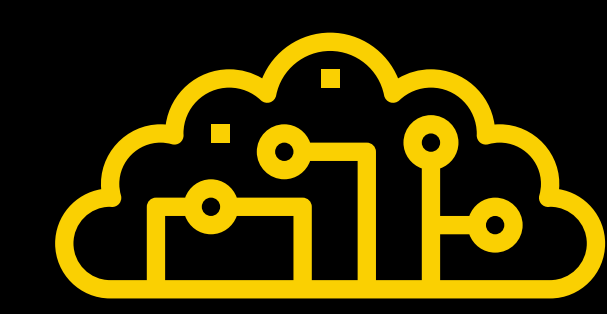


So can we conclude MODEL - Y works best in generating more revenue???

NO!

Running an A/B test is pretty simple.
However, you might run into a few snags
with A/B testing.

This method commonly performed
using a traditional statistical inference
approach grounded in a hypothesis test



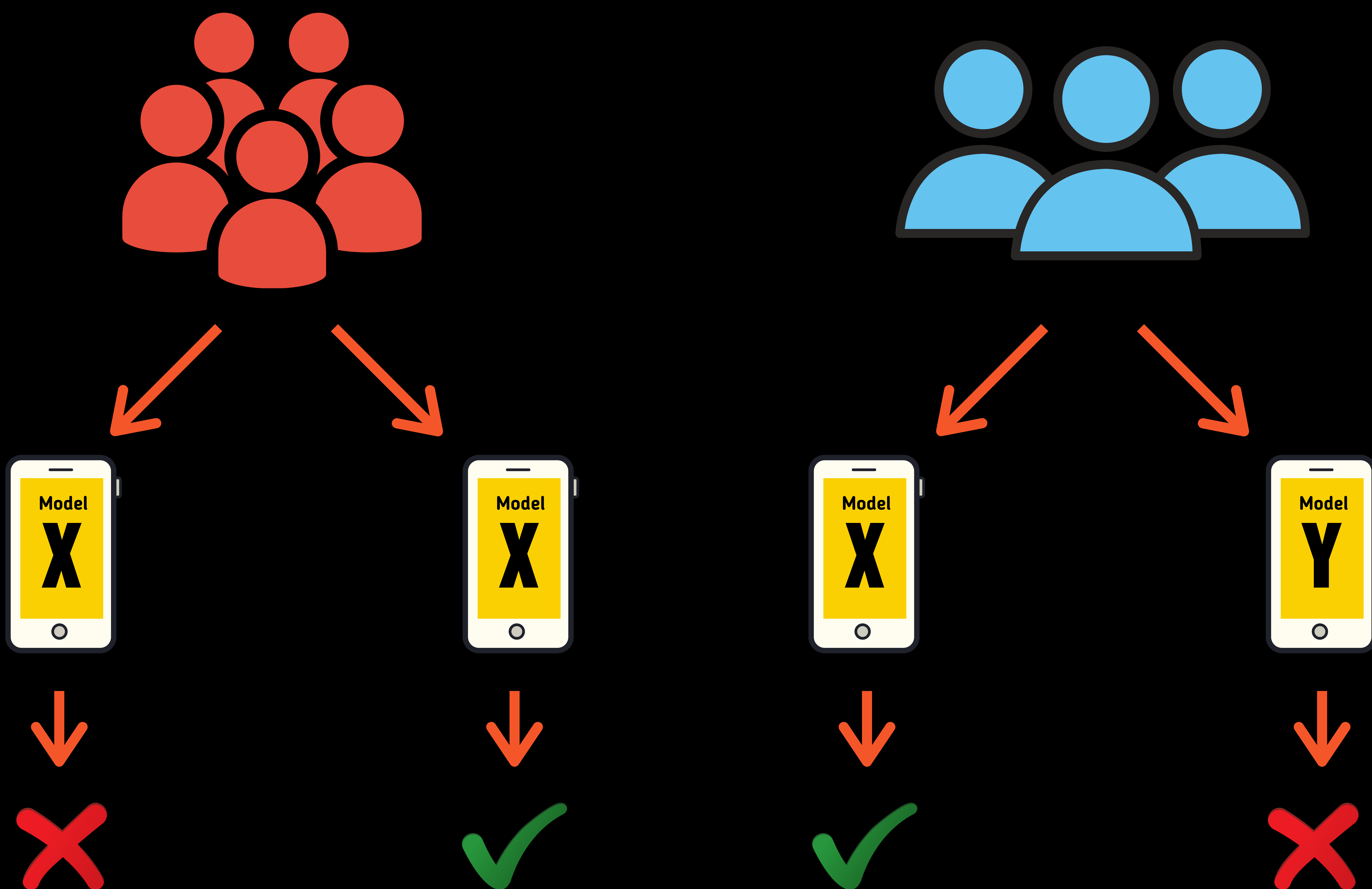
This traditional method has one problem where it can only consider one factor to analyse which one works best.

But the world has many factors like visiting users can have different characteristics like age, gender etc., the amount of time they spend, redirected by a different source etc....

So we cannot do an A/B test of five different models considering a single feature each time.

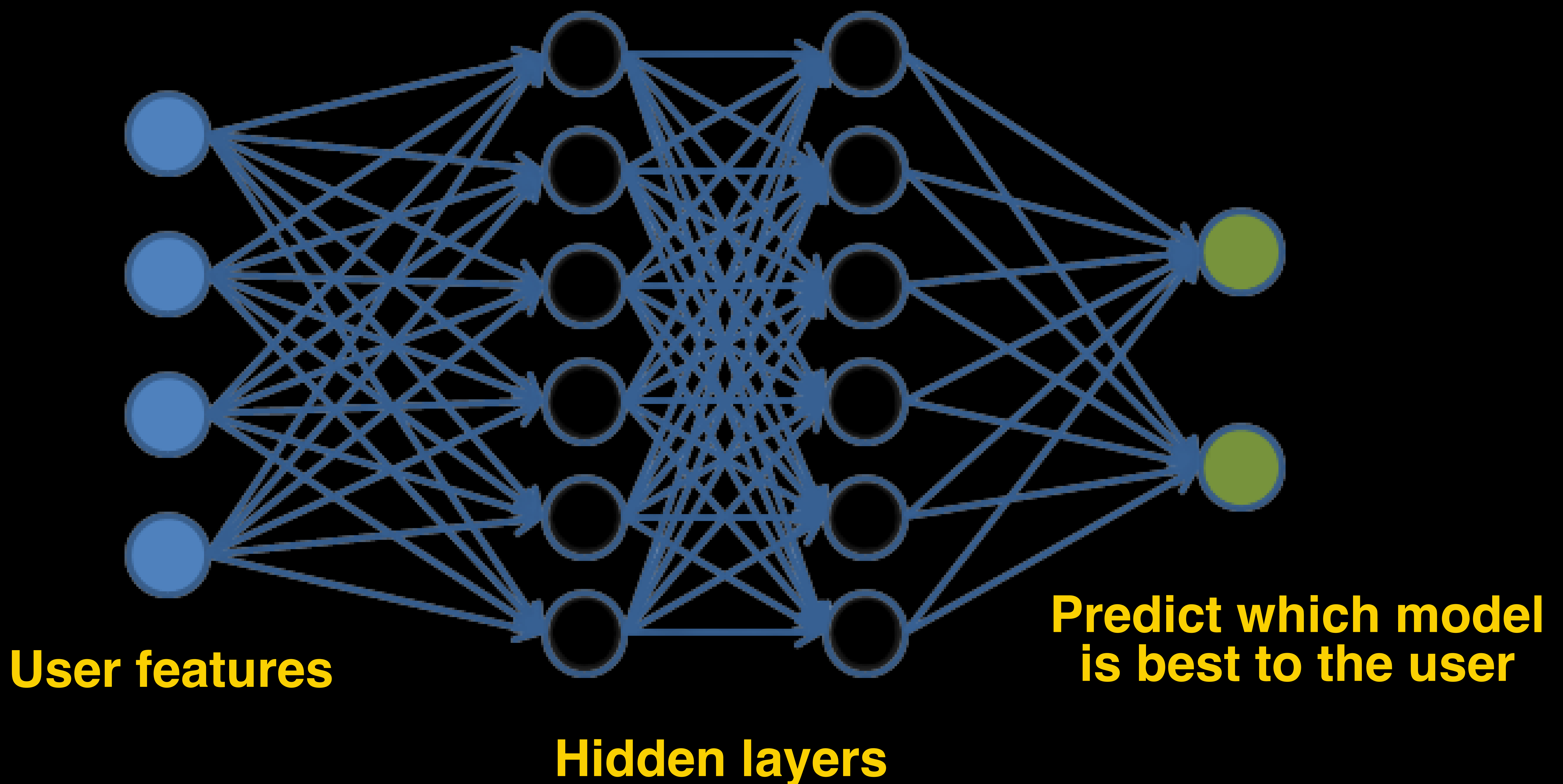
Another Problem





There may be people who still like Model - X then Y in group B and Few people in group A like the model - X and few don't.

If we decide to go with MODEL Y, you might end up alienating people who prefer MODEL X. thereby hurting your overall Conversion Rate.



To solve the above problems we can use machine learning to optimize the A/B testing by considering all the features but still, the A/B test is a time-consuming process.

Instead, we can only use ML to consider all the factors and predict which model is best for the user and provide that to increase the revenue.



Which one do you choose to test your ML models in production?

