

# Churn Prediction Modeling and ML System Design

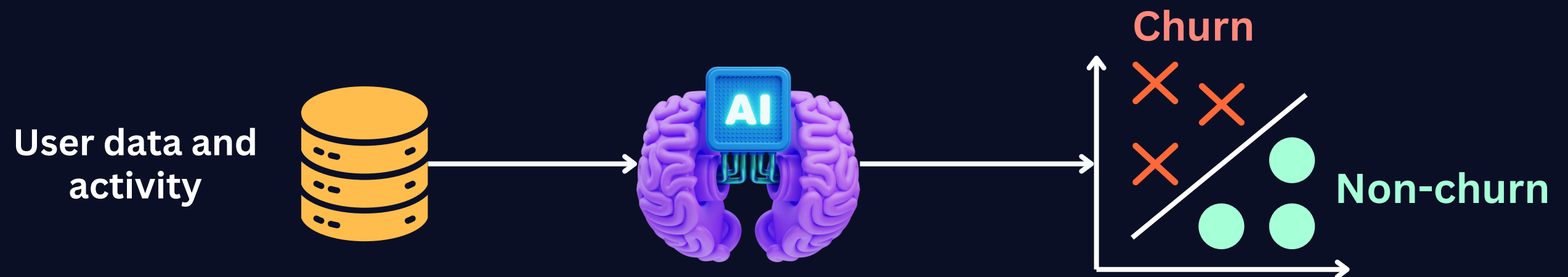
## Business Problem

Churn prediction systems identify users or customers likely to cancel a subscription, stop purchasing, or disengage from a service.

Common business goals include:

- Improving customer retention by identifying high-risk users early.
- Reducing churn-driven revenue loss.
- Supporting targeted interventions such as discounts, reactivation offers, or proactive outreach.

Churn prediction can be modeled as binary classification (will churn / will not churn) or as risk scoring (likelihood to churn).



# Churn Prediction ML System Design

## Typical Features:

- **Recency:** Days since last login or activity
- **Frequency:** Number of logins, sessions, or purchases in the last X days
- **Engagement duration:** Average session length over the last week or month
- **Feature usage counts:** How often key features were used (e.g., reports generated, messages sent)
- **Plan or tier info:** Subscription type, feature access
- **Support activity:** Number of support tickets filed, time to resolution
- **Billing patterns:** Failed payments, payment method changes, recent upgrades/downgrades
- **Marketing interaction:** Click-through rate on emails, offers redeemed
- **Net promoter score (NPS)** or survey feedback when available
- **Account age:** Days since sign-up or subscription
- **Inactivity streak:** Longest stretch of inactivity in recent time window

## Typical Target:

- Churn / non-churn label within churn window (e.g. next 14 days)

## Typical Models:

- Random Forest and Gradient Boosting (XGBoost, LightGBM)
- Logistic Regression for interpretable binary classification
- Neural Networks for behavioral and event sequence modeling
- Autoencoders or Isolation Forests for anomaly-based churn risk

# Churn Prediction ML System Design

