Relax Challenge Summary

Adam Reichenbach
Springboard Data Science Career Track June 2024 Cohort
LinkedIn: https://www.linkedin.com/in/adam-reichenbach/
GitHub project Repo: https://github.com/DJSydon

Objective

Identifying which factors can predict user adoption. A **user** is an individual who logged into the product on three separate days within at least one seven-day period.

Approach

We labeled users based on login data and merged with the metadata for 12,000 users. Feature engineering included signup method, invite status, and engagement duration. A logistic regression model was trained for interpretability.

Key Findings

- For signups: **Google authentication sign ups** had the highest adoption rate, followed by regular **signups**, suggesting that users who opt in themselves are more likely to engage.
- **Guest invites** performed surprisingly well, suggesting short bursts of focused activity.
- **Longer time since signup** strongly correlates with adoption.
- **Invited** users were more likely to be adopted than non-invited users, but still less than self_signups.
- **Org-level effects** (via org_id) may play a role in adoption clustering.

Considerations
- While a logistic regression model was chosen for interpretability; non-linear models like random forests may uncover deeper interactions.
- Missing login timestamps were imputed with the median; more robust time-aware methods could improve accuracy.

Further Exploration
- Analyze adoption by inviter or organization groupings.
- Incorporate product usage beyond logins.
- Explore retention metrics beyond initial adoption.

Adoption Rate by Creation Source

Distribution of org_id by Adoption Status

Distribution of creation_year by Adoption Status

Distribution of creation_month by Adoption Status

Distribution of creation_day by Adoption Status

Distribution of creation_weekday by Adoption Status

Distribution of days_since_creation by Adoption Status