Module 4

Workshop 1: Feature Engineering Fundamentals in Practice

Find the link!

What do these images have in common, and how does this relate to feature engineering?

Share your thoughts in the chat!

Module 4

Workshop 1: Feature Engineering Fundamentals in Practice

Agenda

- **Review:** Recap key concepts.
- **Practice:** Feature engineering for churn prediction.
- **Closing:** Key takeaways and next steps.

Learning objectives

- **Identify basic data quality issues** in a real-world dataset and apply appropriate data cleaning techniques to address them.
- **Apply encoding techniques** to convert categorical variables into model-ready features.
- **Create new features** that represent customer tenure segments, service usage levels, and spend patterns to capture domain-specific insights.
- **Evaluate the impact of engineered features** on machine learning model performance and interpretability.

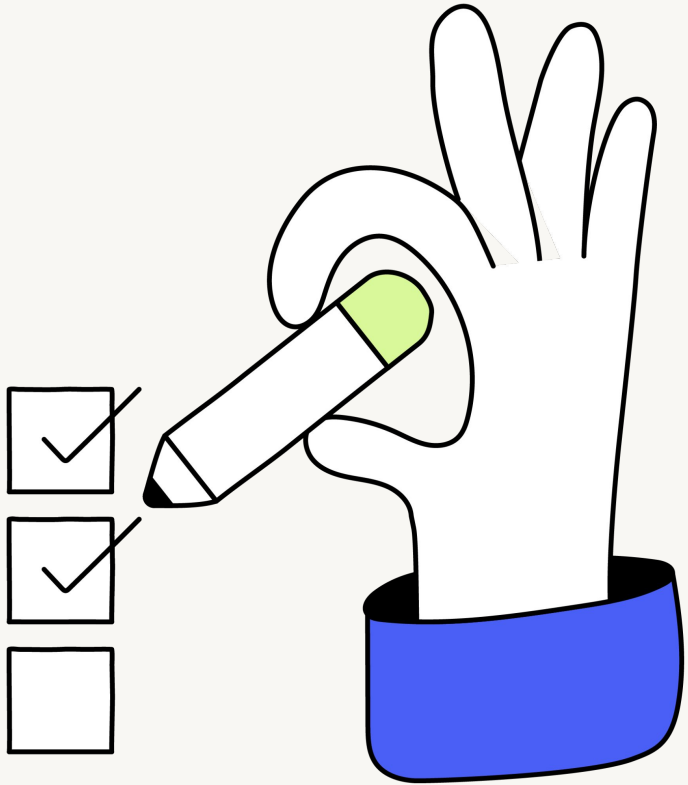
Async review: Unit 1

Go to My Multiverse, Module 4
Workshop 1

Navigate to the **“Async review”** page

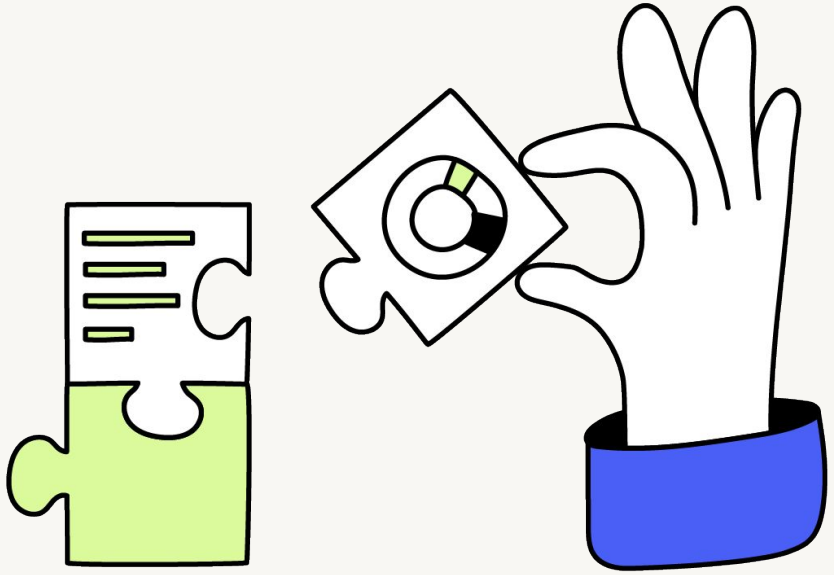
Review the summary of Unit 1, then test your knowledge with the following quiz:

Feature engineering checkpoint



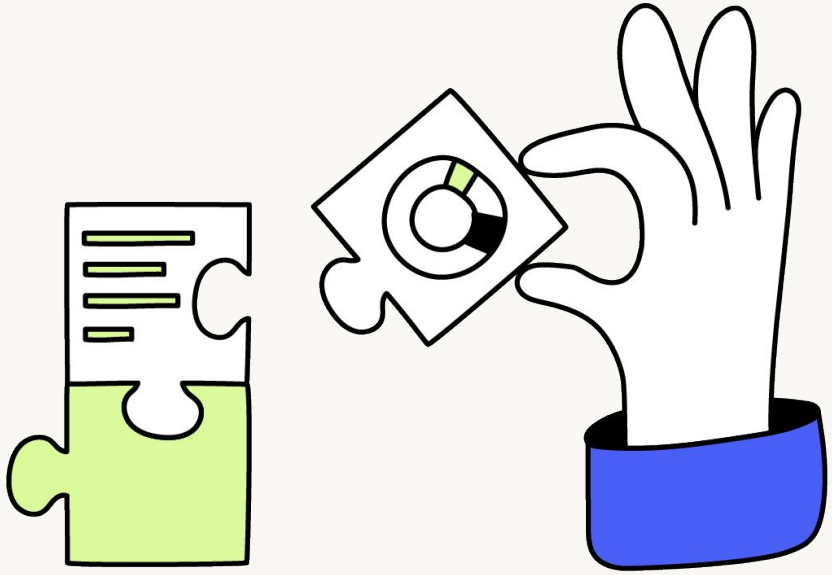
Importance of feature engineering in ML models

- Feature engineering is essential for improving model performance, interpretability, and robustness.
- The quality of engineered features directly affects a model's ability to learn patterns from data.



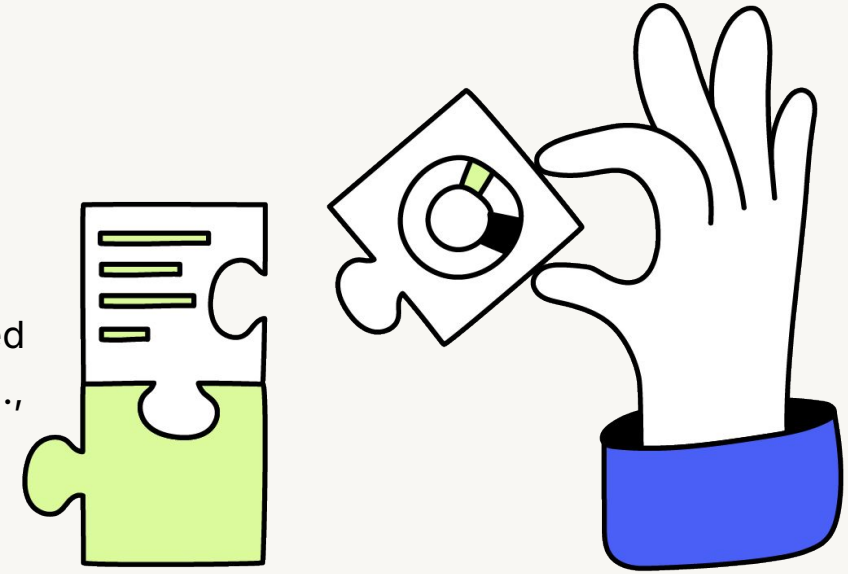
Variable types

- **Continuous:** Numeric values with infinite possibilities (e.g., tenure in months).
- **Discrete:** Numeric values with finite, countable steps (e.g., number of support calls).
- **Categorical:** Distinct groups without order (e.g., contract type).
- **Ordinal:** Ordered categories (e.g., satisfaction: low, medium, high).



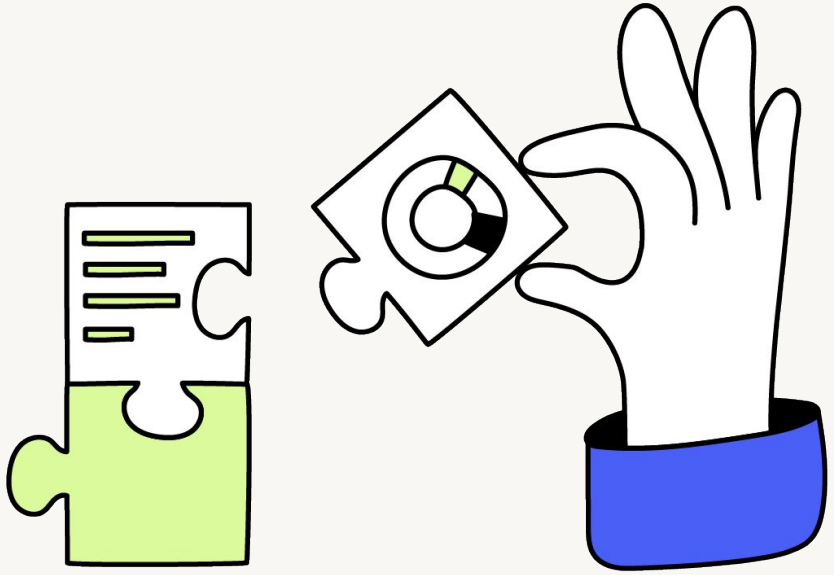
Feature engineering techniques

- **Data cleaning:** Fixing missing values, inconsistencies, and formatting issues.
- **Encoding:** Converting categorical data into numerical form (e.g., one-hot, ordinal).
- **Scaling:** Standardizing numerical features to ensure consistency.
- **Feature selection:** Removing irrelevant or redundant features for better performance.
- **Feature synthesis:** Creating new features based on domain knowledge and predictive value (e.g., inferring gender from names if relevant).



Feature engineering and model performance

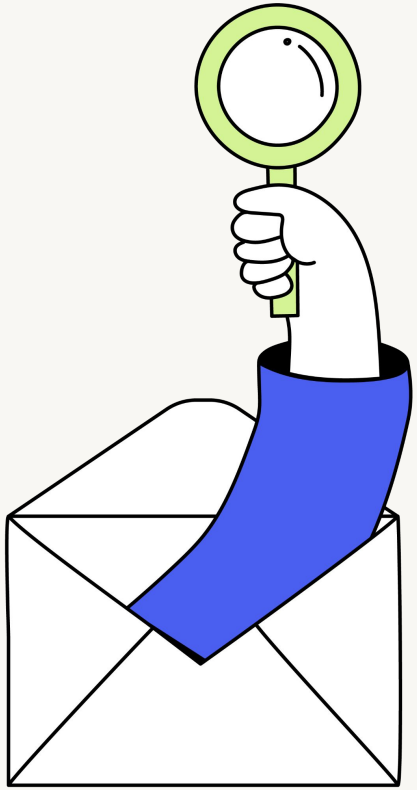
- Well-engineered features **boost accuracy** and prevent misleading correlations.
- Feature selection **reduces overfitting** and improves generalization.
- Transformations like log scaling, binning, and domain-specific features **uncover hidden patterns** for better predictions.



Skills application: Feature engineering for churn prediction

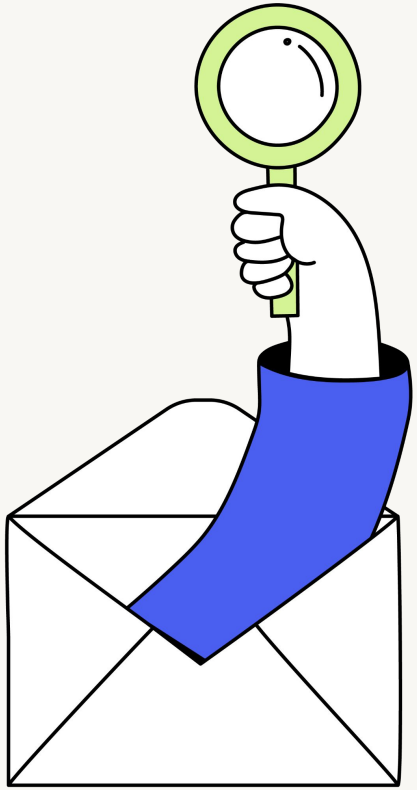
Skills application introduction

- **What you'll do:**
  - Apply data preprocessing techniques to clean and structure the dataset.
  - Engineer new features that capture customer behavior and improve model performance.
  - Explore interaction features and aggregated metrics to refine predictions.



Skills application instructions

- Work in breakout rooms in small groups.
- Use the Jupyter Notebook to work on applying feature engineering techniques to the Telco Customer Churn dataset
- The challenge is designed to be solved individually, but collaboration is encouraged.
- Find all instructions and resources on Ariel.



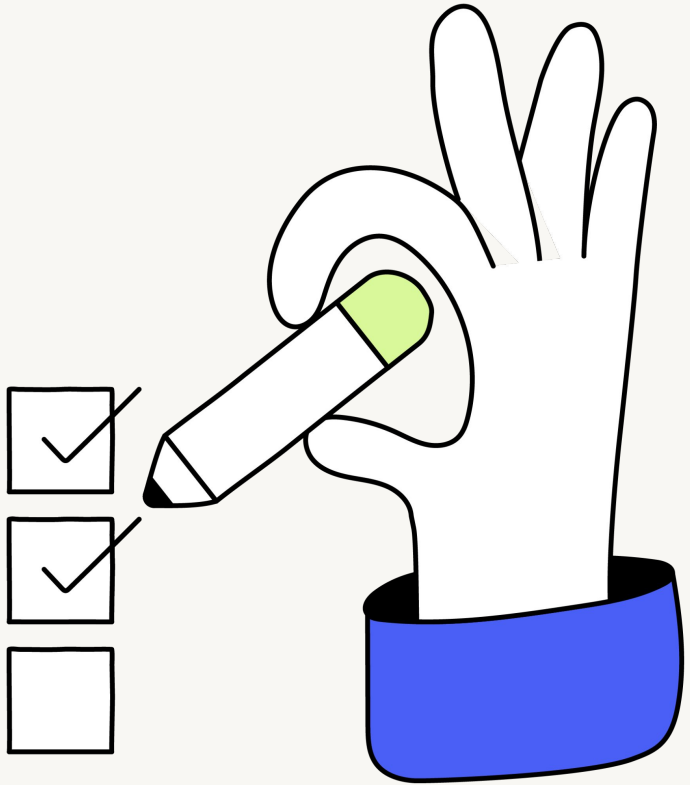
Access the skills application

Go to My Multiverse, Module 4 Workshop 1

Navigate to the **“Skills application”** page

Start working on the activity challenges

Be prepared to share in 25 minutes when we come back together



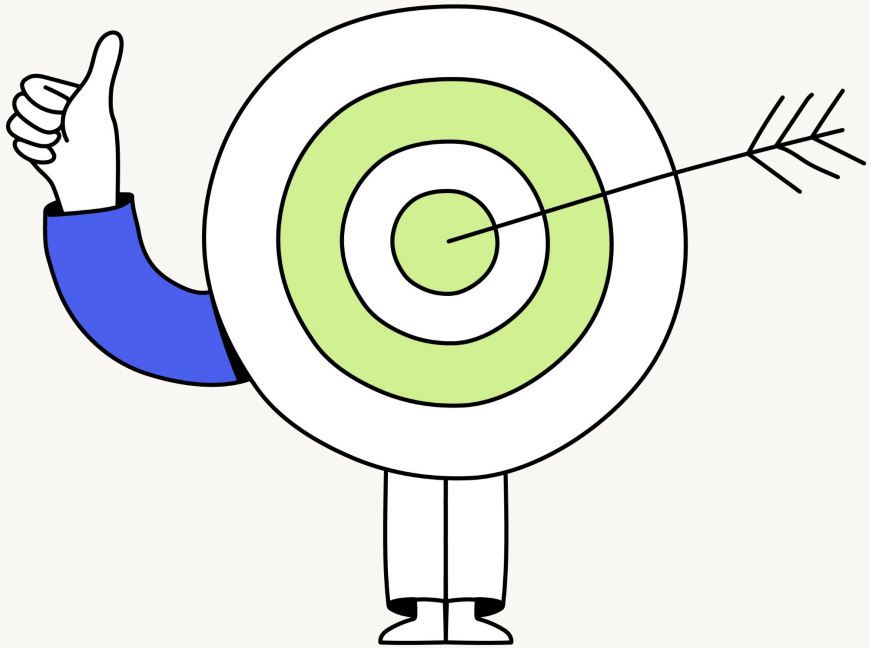
Skills application debrief

Skills application debrief

What feature engineering techniques did you apply, and how did they impact the model’s ability to predict customer churn?

Key takeaways

1. Feature engineering drives better models.
2. Domain knowledge is key.
3. Engineered features enhance business interpretability.
4. Reducing churn via proactive retention.



Applying key takeaways to your organization

Applying key takeaways

1. What is a key outcome or behavior your organization wants to predict?

2. What features could you engineer that would likely signal this outcome in advance?

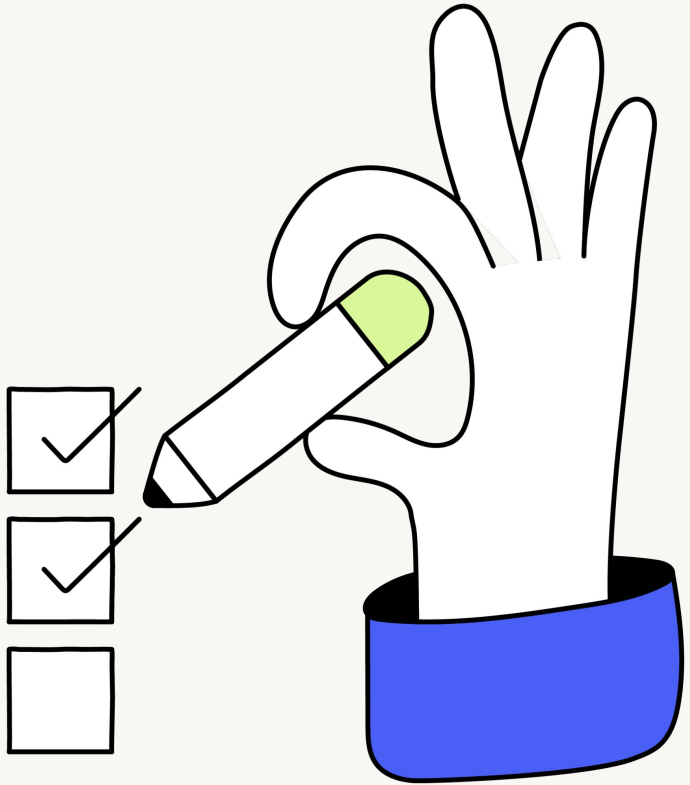
Share your plan to apply new skills

Go to My Multiverse, Module 4
Workshop 1

Navigate to the discussion activity page **“Apply new skills to your role”**

Take 5 minutes to share how you will apply this skill in your role

Take time this week to read what others have written and get new ideas!



Closing out

- **Next up:** Advanced Feature Engineering for Time-Series Data
- **If you have questions, you can:**
  - Chat with Atlas
  - Reach out to your Cohort Coach
  - Sign up for Tutoring
- **Log your OTJ hours**
- **Take the Workshop Survey** on My Multiverse

multiverse