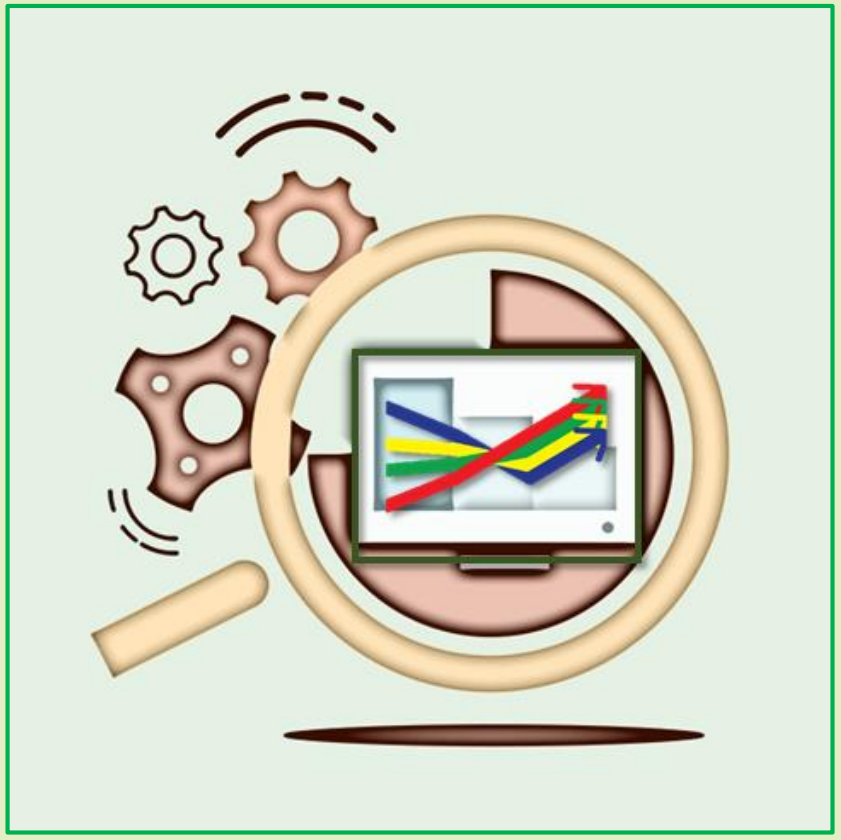
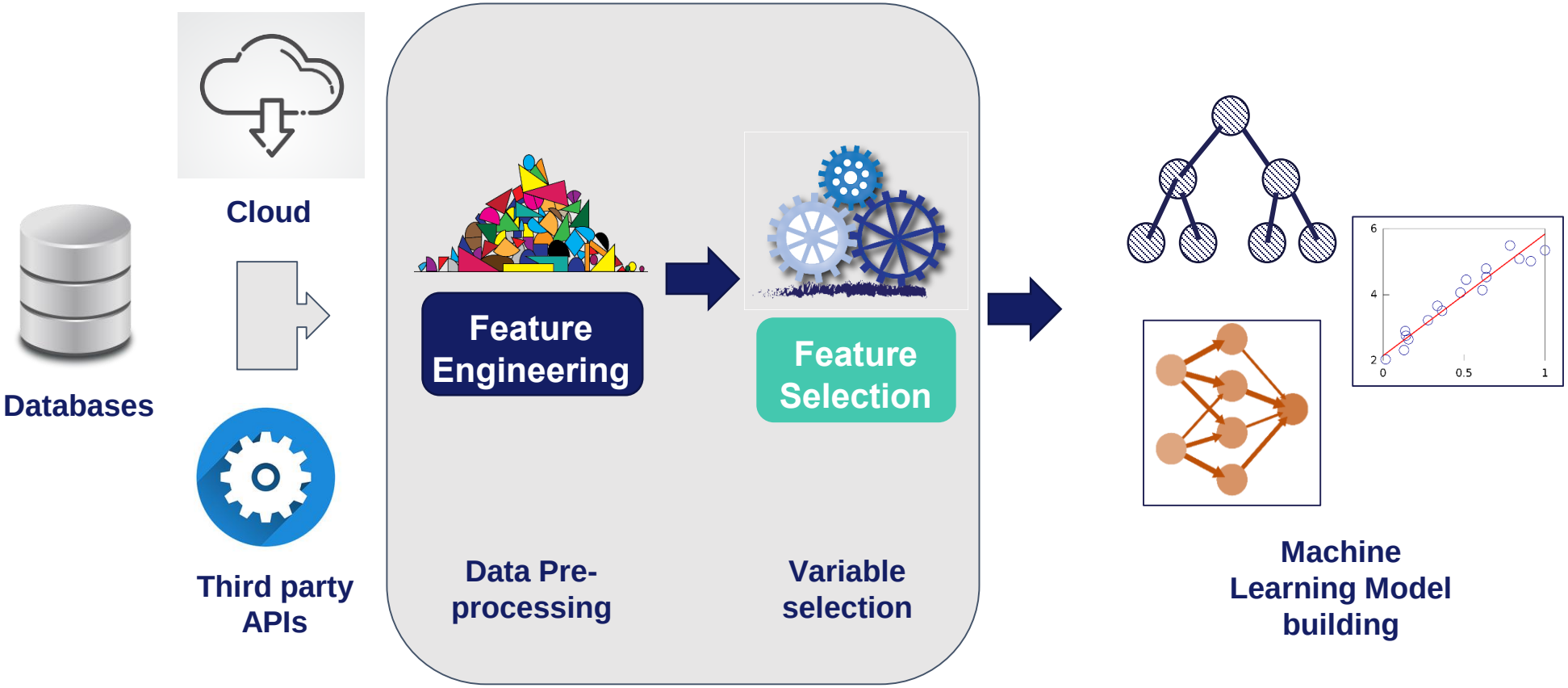
Train In Data

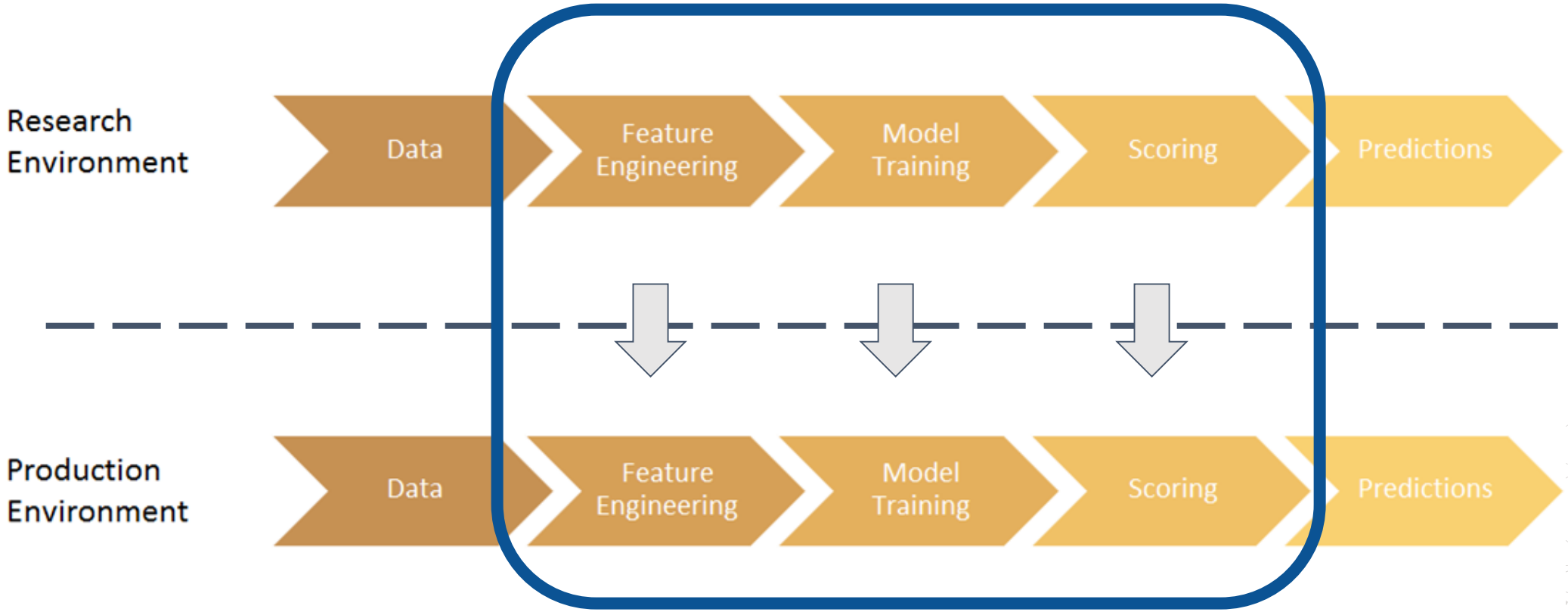


Research Environment - Creating a ML Model

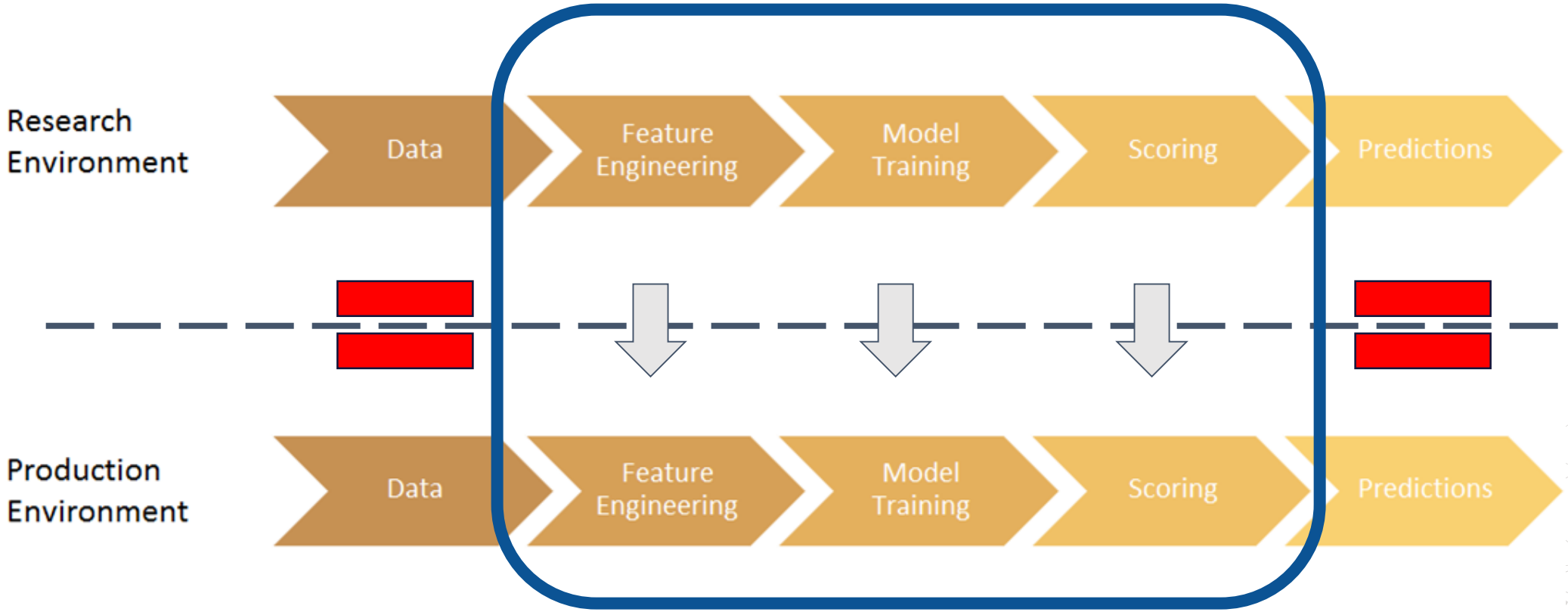
Machine Learning Pipeline



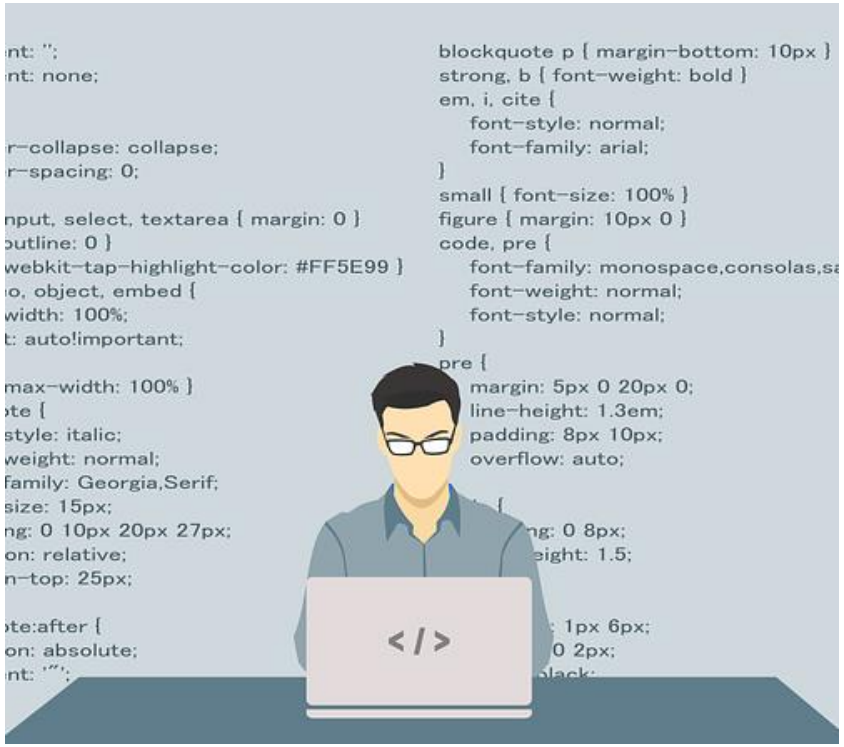
Deployment of ML Pipeline



Reproducibility



Principles of ML Deployment



- Challenges of traditional software
  - Reliability
  - Reusability
  - Maintainability
  - Flexibility
- Additional challenges specific to Machine Learning
  - Reproducibility

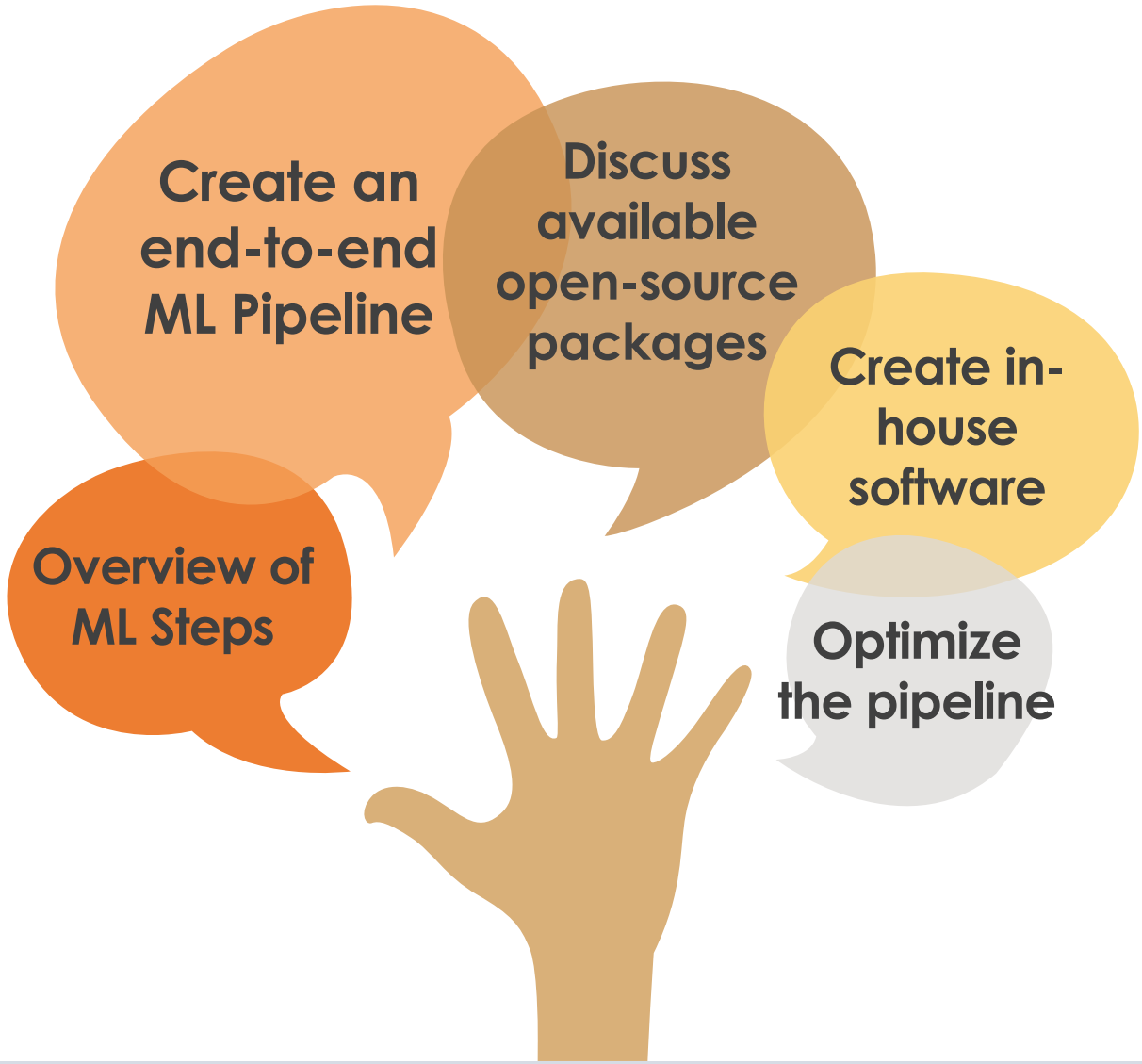
Open-source

- Increase Performance
  - ✓ Ready to use code
  - ✓ Off-the-shelf algorithms
- Maximise reproducibility
  - ✓ Versioning
- Reliability
  - ✓ Testing
- Reusability, maintainability
- Minimise deployment times

In-house software

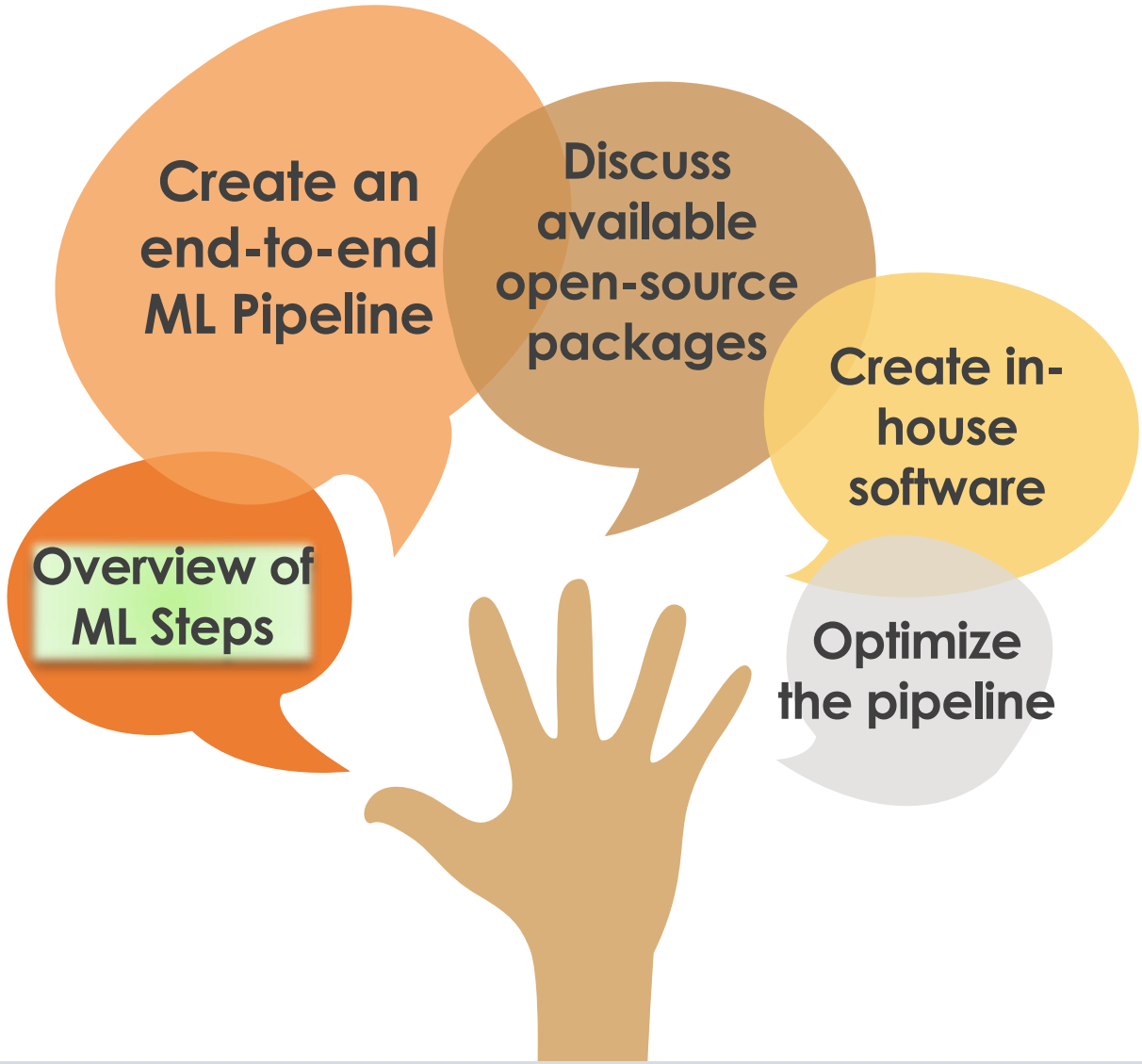
- Versioned
  - ✓ Reproducibility
- Tested
  - ✓ Reliability
- Shareable
  - ✓ Reusability
- Minimise deployment times

5 STEPS



1. Overview of ML steps
   - Feature engineering and selection
   - Model building
2. Create a ML Pipeline
3. Available Open-Source
   - Scikit-learn, Feature-engine, Category encoders
4. Create in house software
   - Use of the scikit-learn API
5. Re-create an optimized pipeline
   - Python environments

5 STEPS



1. **Overview of ML steps**
   - **Feature engineering and selection**
   - **Model building**
2. Create a ML Pipeline
3. Available Open-Source
   - Scikit-learn, Feature-engine, Category encoders
4. Create in house software
   - Use of the scikit-learn API
5. Re-create an optimized pipeline
   - Python environments

Train In Data

THANK YOU

www.trainindata.com