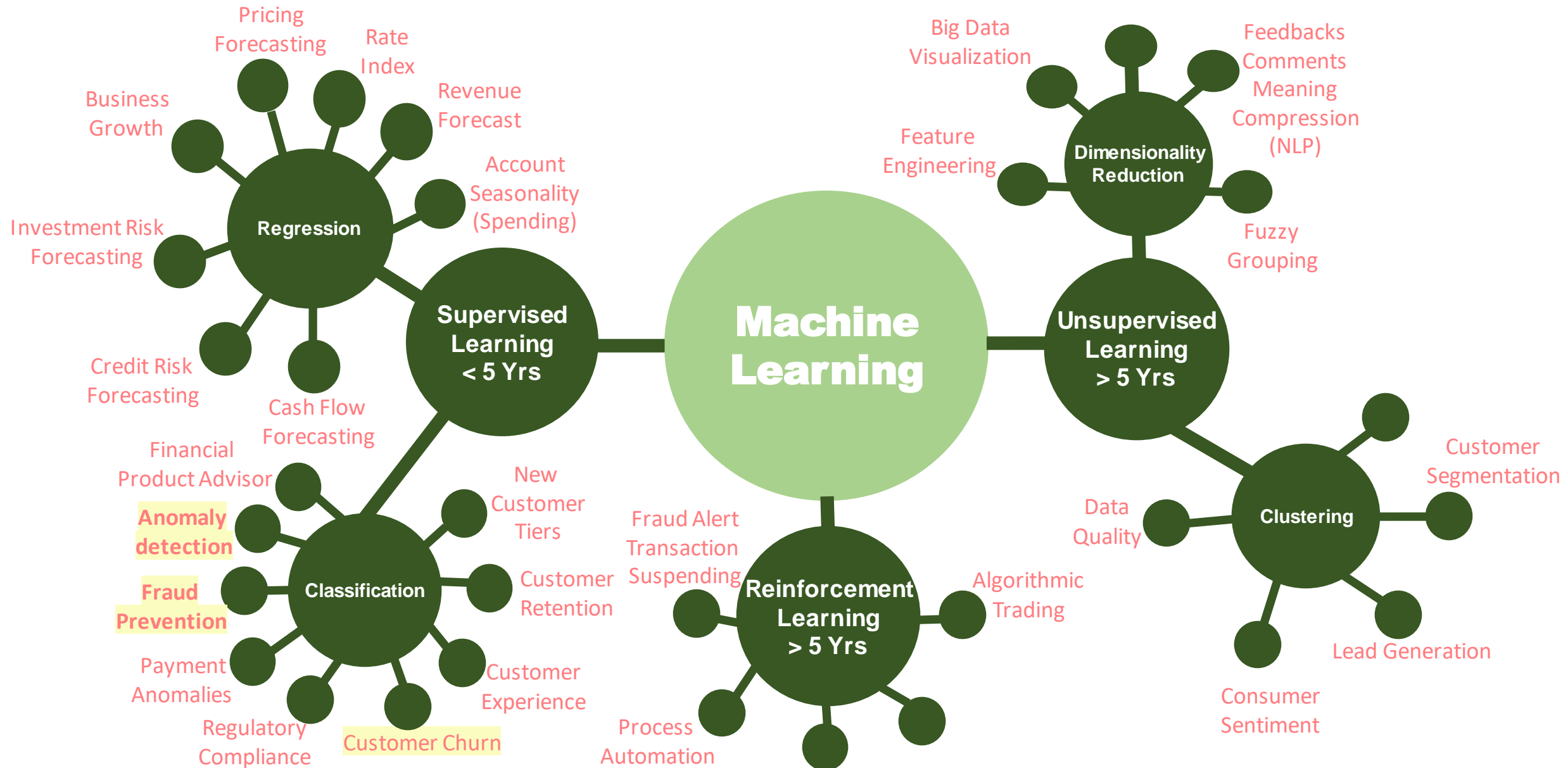


Topics to Discuss

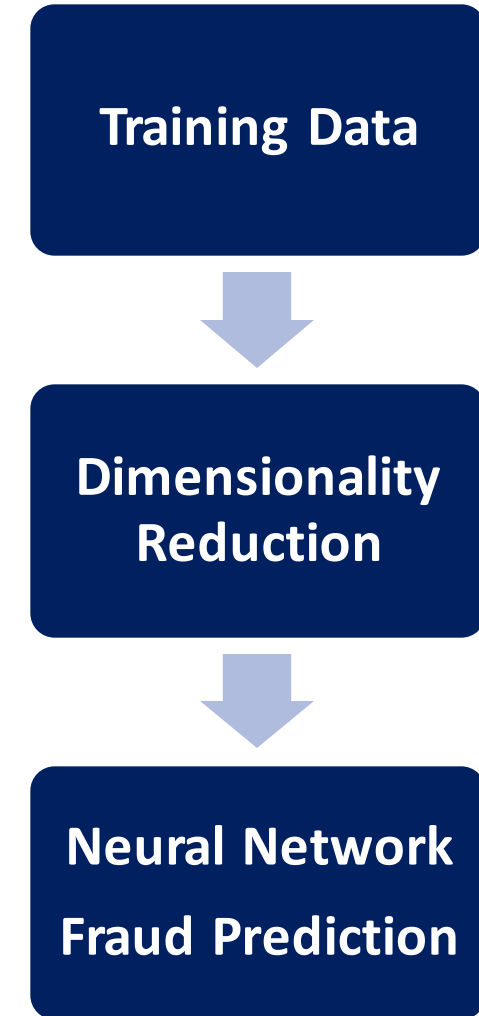
- ML pipeline from POC to Production
- Bring the current ML models to AWS,
 - Data Pipeline
 - Data Models (Training/Testing/Forecast Data) & ML Database
 - Gitlab Python Code Control & (Jupyter Notebook)
 - Kubeflow/Dataiku deployment
 - ML Data consumption (PostgreSQL/AWS Redshift) :Tableau Visualization & MLaaS API
 - ML Wiki
- ML K8 & SparkSQL Architecture Scalability
- Data Compliance & Strategy
 - Coordinate between Data Centers on AWS
 - Data Masking
- ML Platform:
 - Platform Tools: SageMaker/Kubeflow/H2O/Data Robot/Dataiku
 - Dataiku to orchestrate(CI/CD/CT) the pipeline flow of ML projects
 - Templates for different project size
- ML Projects:
 - See Next Slide
- ML Roadmap: Showcase & ML workshops to communicate with businesses, Peers & Leaders
- ML Leadership: Communication, Problem Solving Skills, Mentor/Guideline/Goal with team

TAXONOMY OF MACHINE LEARNING METHODOLOGIES IN BANKING

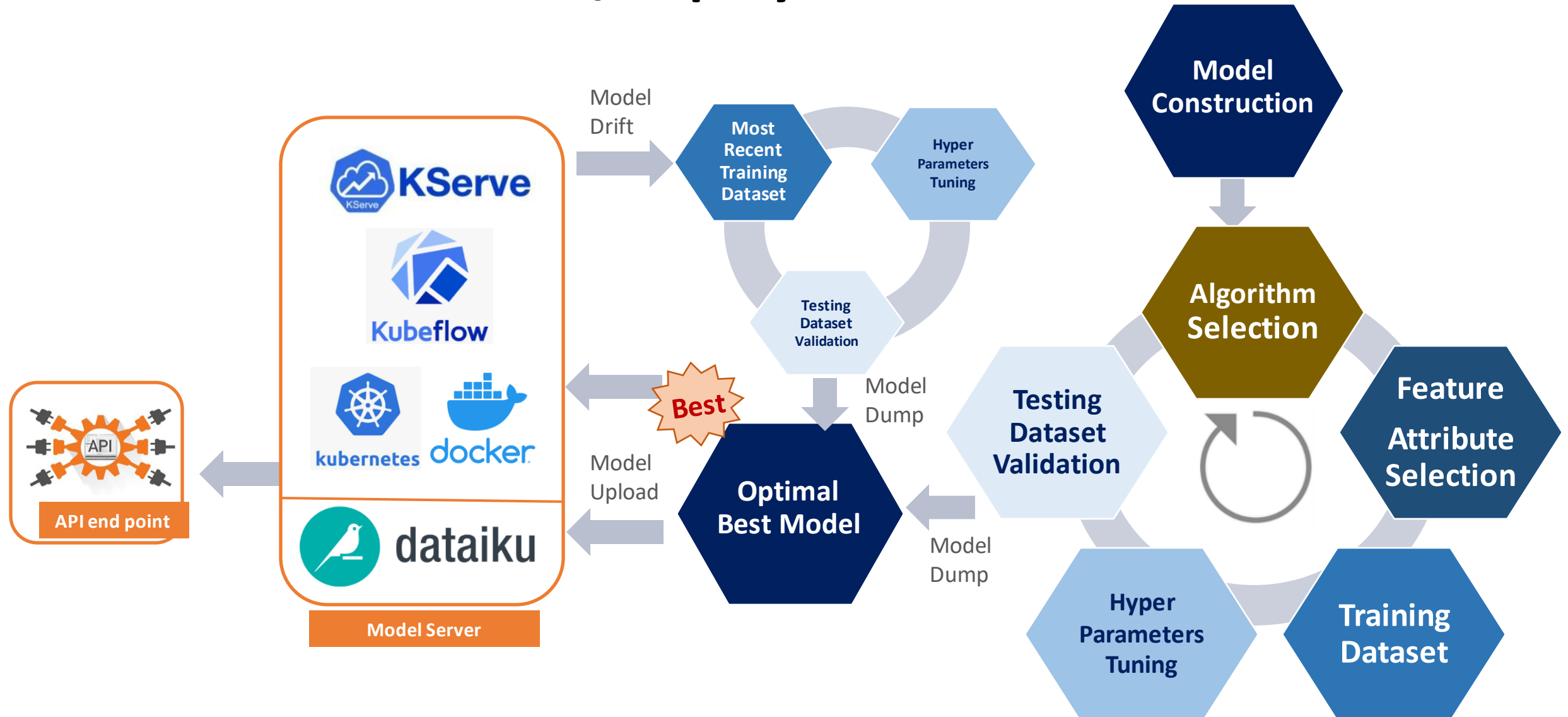


Fraud Prediction Risk Score

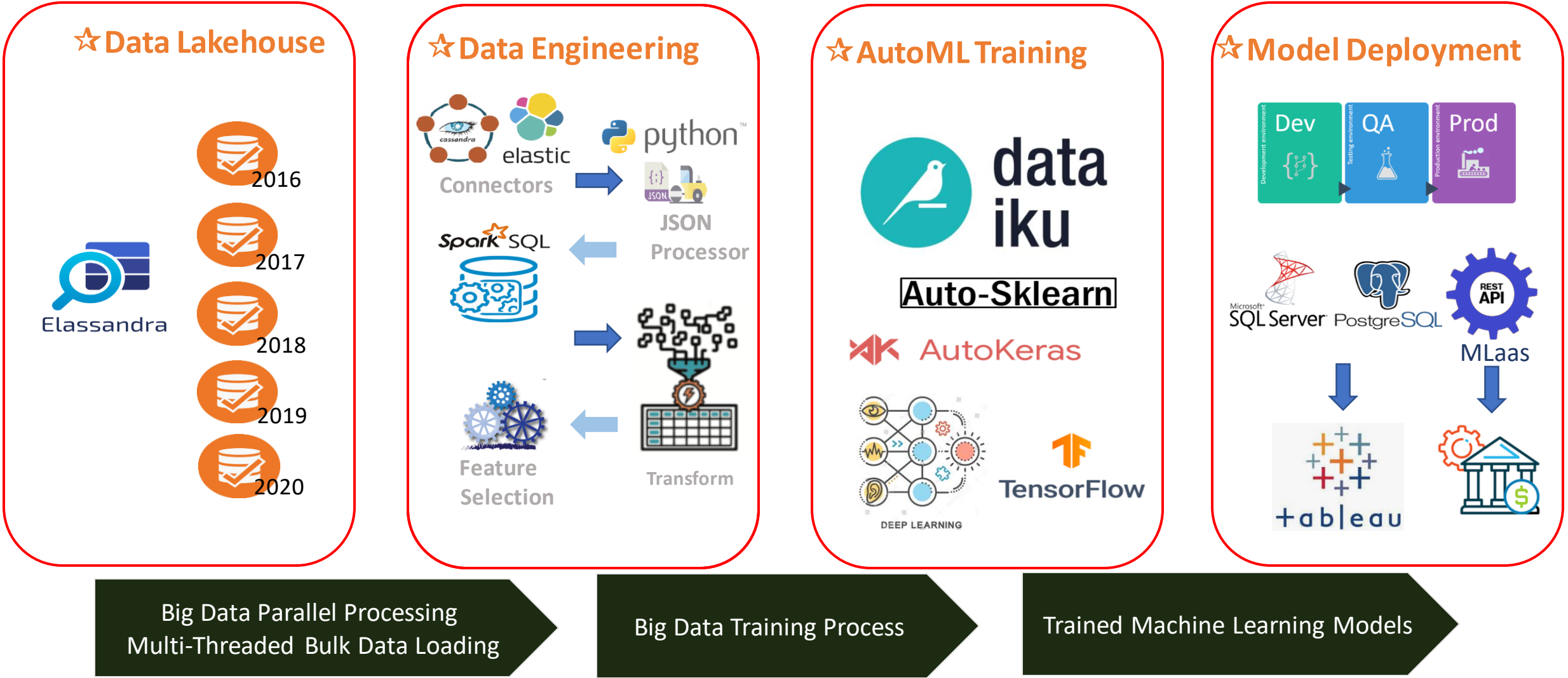
- **Features include:**
 - Transaction Type
 - Personal Information(Age, Race, Education, etc)
 - Data (\$\$, debt, credit)
 - IP address
 - Location
 - Credit Score
 - Payment History Score
 - Current Income
 - Credit History
- **Measures:**
 - **Fraud Risk Score**
- **Date:**
 - N/A
- **Frequency:**
 - N/A



ML Model Validation/Deployment



Big Data Machine Learning Data Architecture



Neural Network

Loss
Function

Backward Propagation

Gradient Descent & Learning Rate

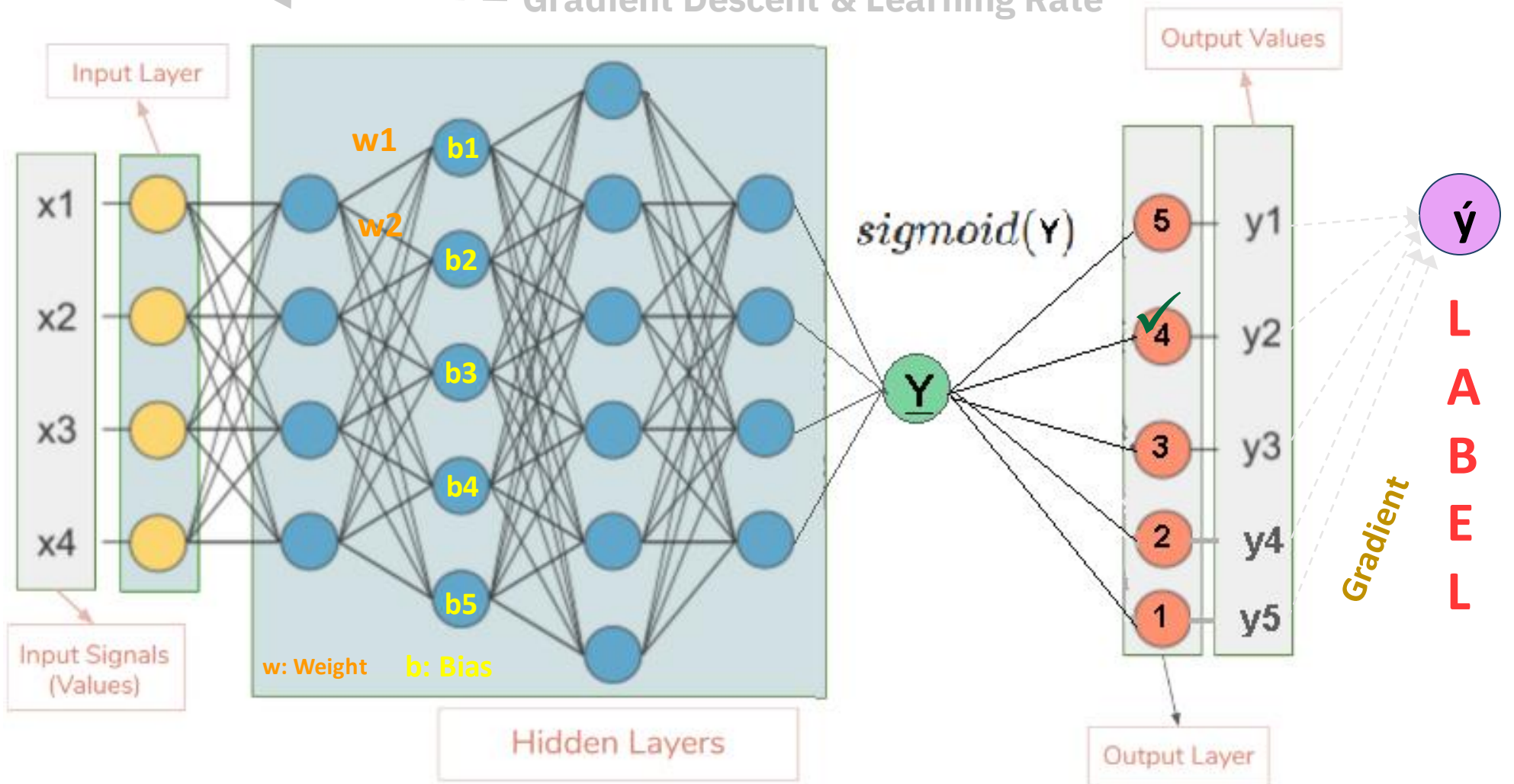
Features

Credit
Score

Location

IP Risk

Amount
Range



Enterprise **ML** Platform Project Management

