**Machine Learning System Design**

**Table of Content**

1. Introduction to Machine Learning Systems Design
2. Data Engineering Fundamentals
3. Training Data
4. Feature Engineering
5. Model Development and Offline Evaluation
6. Model Deployment and Prediction Service
7. Data Distribution Shifts and Monitoring
8. Continual Learning and Test in Production
9. Infrastructure and Tooling for MLOps

**Introduction to Machine Learning Systems Design**

* Business and ML Objectives
* Requirements for ML Systems
* Iterative Process
* Types of ML Tasks
* Classification versus regression
* Objective Functions
* Mind Versus Data

**Data Engineering Fundamentals**

* Type of different source
* Third-party data collector
* Comparison of data-serialization formats
* JSON
* No-SQL
* Structured data vs Unstructured data
* Transactional and Analytical Processing
* ETL: Extract, Transform, and Load
* REST and RESTful

**Training Data**

* Sampling
* Labeling
* Class imbalance
* Challenge with imbalanced data
* Handling Class Imbalance
* How to modify Loss Function
* Data Augmentation

**Feature Engineering**

* Handling Missing Values
* Scaling
* Discretization
* Encoding Categorical Features
* Feature Crossing
* Discrete and Continuous Positional Embeddings
* Data Leakage
* Common Causes for Data Leakage
* Engineering Good Features

**Model Development and Offline Evaluation**

* Six tips for model selection
* Ensembles
* Boosting
* Stacking
* Experiment tracking
* Versioning
* Debugging ML
* ML model to fail
* Some debugging techniques
* Data parallelism
* Model parallelism
* **AutoML**

1. Soft AutoML: Hyperparameter tuning
2. Hard AutoML: Architecture search and learned optimizer

* FOUR PHASES OF ML MODEL DEVELOPMENT
* Model Offline Evaluation
* Evaluation Methods

**Model Deployment and Prediction Service**

* Machine Learning Deployment Myths
* Batch Prediction Versus Online Prediction
* Unifying Batch Pipeline and Streaming Pipeline
* Model Compression
* ML on the Cloud and on the Edge

**Data Distribution Shifts and Monitoring**

* Software System Failures
* ML-Specific Failures
* Edge cases
* Data Distribution Shifts
* Detecting Data Distribution Shifts
* Monitoring and Observability

**Continual Learning and Test in Production**

* continual learning
* Stateless Retraining Versus Stateful Training
* Continual Learning Challenges
* Four Stages of Continual Learning
* How Often to Update Your Models
* Test in Production

**Infrastructure and Tooling for MLOps**

* Storage and Compute
* Development Environment
* Cron, Schedulers, and Orchestrators
* To help with debugging and maintenance
* Feature Store

------------------------------------------------------------------