Program Structure

1st Semester

Mathematical Foundations For Machine Learning

- Vectors, norms, distances, matrix operations (determinant, inverse, orthogonal)
- Rank, span, bases, eigenvalues, and eigenvectors, SVD, PCA and image compression
- Probability: discrete and continuous distributions, Bayes' theorem, expectation, covariance, law of large numbers, CLT, MLE, Bayesian estimation, hypothesis testing, CIs
- Multivariable calculus: gradients, Jacobian, Hessian, chain rule
- Optimization: critical points, convexity, gradient descent variants, automatic differentiation, and backpropagation
- Python and OOP: classes, inheritance, polymorphism, unit tests
- Graphs and trees: adjacency, DFS and BFS, Dijkstra, spanning trees, ML uses
- Spatial (2D and 3D): coordinate systems, transformations, KD-trees, ball trees, point cloud basics
- Approximate string matching: Levenshtein distance, n-grams, Jaccard similarity, fuzzy hashing

Machine Learning

- ML basics: supervised, unsupervised, and reinforcement learning, datasets, workflows, history, Python libraries
- Linear and logistic regression, MSE, cross-entropy, gradient descent, Ridge, Lasso
- Evaluation: train, validation, and test sets, RMSE, accuracy, precision, recall, F1, ROC-AUC, cross-validation, overfitting remedies
- Algorithms: k-NN, Naïve Bayes, decision trees, ensembles — random forest, gradient boosting
- SVM: maximum margin, hinge loss, kernels
- Introduction to neural networks: perceptron, MLP, backpropagation concept, model selection tips, Expert Weekend
- Hyperparameter tuning: grid search, random search, Bayesian optimization, bias–variance tradeoff, feature selection, Kaggle-style case
- Clustering: k-means, hierarchical, DBSCAN, inertia, silhouette score, segmentation, compression
- Dimensionality reduction: PCA, t-SNE, UMAP
- Association rules: Apriori algorithm, market basket analysis

- Particular topics: graphical models, recommender systems, time series, RL link, ethics
- Neural nets: neurons, weights, biases, activations, forward pass, cross-entropy loss
- Training neural networks: backpropagation, SGD, learning rate, epochs, batches, two-layer demonstration
- Frameworks: PyTorch and TensorFlow graphs, autograd, NumPy scratch build

2nd Semester

Deep Learning And Applications

- Neural network basics: perceptron, ANN layers, weights, biases, activations (sigmoid, tanh, ReLU), forward pass, cross-entropy loss
- Backpropagation and training: gradients via chain rule, SGD, learning rate, epochs and batches, two-layer XOR demonstration, NumPy scratch versus PyTorch and TensorFlow autograd
- Deep network training: vanishing and exploding gradients, Xavier and He initialization, ReLU, regularization — L2, early stopping, Dropout, BatchNorm, optimizers — Momentum, Nesterov, RMSProp, Adam, LR schedules; MNIST experiment
- CNNs: convolutional filters, stride, padding, and pooling, architectures — LeNet, AlexNet, VGG, ResNet, DenseNet, EfficientNet; vision tasks — classification, R-CNN detection, U-Net segmentation, CIFAR-10 lab
- RNNs: hidden state, vanishing gradients, LSTM, GRU gates, applications — sentiment, language modeling, speech; LSTM text-gen lab
- Attention and transformers: soft and self-attention, multi-head, positional encodings, Transformer versus RNN, BERT, GPT, Hugging Face inference demo
- LLMs: GPT-3 and GPT-4 scale, pretraining and fine-tuning, few-shot learning, prompt engineering, foundation models, ethics — bias, misuse; API case study seminar

Introduction To Reinforcement Learning

- RL basics: agent–environment loop, rewards, episodic versus continuing tasks, MDP components — states, actions, reward function, transition probabilities, discount factor γ, objective — maximize discounted return, policy π(a|s), Gridworld and CartPole illustrations
- Value functions: $V^{\pi}(s)$ and $Q^{\pi}(s,a)$, Bellman expectation and optimality concepts
- Dynamic programming with full model: policy evaluation and iteration, value iteration, Gridworld convergence demo
- Model-free TD control: Q-Learning, SARSA, bootstrapping, exploration versus exploitation using ε-greedy strategy

- Policy gradients: parameterized policies, REINFORCE, high variance and actor-critic methods
- Function approximation: neural nets for value and policy (DQN Atari success), overview — model-based versus model-free, DDPG, PPO
- Applications and challenges: robotics, finance, recommender systems, sample efficiency, safety, reward shaping, OpenAI Gym showcase
- Connections: RL as supervised regression for value nets, synergy with deep learning, next steps — Sutton and Barto, advanced courses

Introduction To MLOps (Machine Learning Operations)

- MLOps versus DevOps, ML lifecycle touchpoints, reproducibility — track code, data, and hyperparameters with MLflow
- Git collaboration, data and model versioning via DVC and Git LFS, full checkpoint traceability
- Deployment: model serialization (Pickle or ONNX), REST APIs (Flask or FastAPI), Docker containers, CPU versus GPU, real-time versus batch serving
- CI and CD: unit and integration tests, GitHub Actions and Jenkins build-push, infrastructure as code, orchestrators — Kubeflow, Airflow
- Monitoring: latency, throughput, data and concept drift, rolling accuracy, Prometheus and Grafana dashboards, alerts and retraining triggers, governance and audit logging
- Practice: Yandex production case, Kubernetes scaling, end-to-end pipeline demo, feature stores and ML platform trends