**Syllabus for cycle 2:**

This is the syllabus for cycle 2, we provided you with some resources you can explore for other resources also to boost your understanding skills.

**Phase 1: Boosting & Bagging (XGBoost, AdaBoost)**

**Topics Covered:**

* Ensemble learning overview
* Bagging vs Boosting
* AdaBoost
* XGBoost (intro to gradient boosting)

**Goals:**

* Understand the concept of combining weak learners.
* Apply AdaBoost and XGBoost on classification problems.

**Resources:**

* [Ensemble learning](https://scikit-learn.org/stable/modules/ensemble.html)
* [Boosting and bagging](https://www.youtube.com/watch?v=GM3CDQfQ4sw)
* [Ada Boost](https://www.youtube.com/watch?v=LsK-xG1cLYA)
* [Xgboost](https://www.youtube.com/watch?v=OtD8wVaFm6E)

**Phase 2: Support Vector Machines (SVM)**

**Topics Covered:**

* Linear SVMs
* Kernels: RBF, Polynomial
* Hyperparameter tuning (C, gamma)’
* Introduction to PCA for dimensionality reduction

**Goals:**

* Understand margins, kernel trick.
* Use SVM for classification with linear and non-linear data.
* Apply PCA for feature compression

**Resources:**

* [SVM youtube vedio](https://www.youtube.com/watch?v=efR1C6CvhmE)
* [SVM](https://scikit-learn.org/stable/modules/svm.html)
* [Intro to PCA](https://www.youtube.com/watch?v=FgakZw6K1QQ)
* [PCA](https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c/)

**Phase 3: Unsupervised Learning & Dimensionality Reduction**

**Topics Covered:**

* k-Means clustering
* Hierarchical clustering (dendrograms)
* t-SNE & PCA for visualizing high-dimensional data

**Goals:**

* Discover patterns in unlabeled data
* Visualize clusters using t-SNE/PCA

**Resources:**

* Vedios: [kmeans](https://www.youtube.com/watch?v=4b5d3muPQmA) [heirarchial](https://www.youtube.com/watch?v=7xHsRkOdVwo) [t-sne](https://www.youtube.com/watch?v=NEaUSP4YerM)
* [Clustering](https://scikit-learn.org/stable/modules/clustering.html#k-means)
* [t-sne and pca](https://www.datacamp.com/tutorial/introduction-t-sne)

**Phase 4: Matrix Factorization & SVD (TF-IDF, PCA)**

**Topics Covered:**

* TF-IDF explanation & use in NLP
* Singular Value Decomposition (SVD)

**Goals:**

* Extract features from text using TF-IDF
* Use SVD for topic extraction
* Apply PCA for feature compression

**Resources:**

* [Tf-idf](https://scikit-learn.org/stable/modules/feature_extraction.html#text-feature-extraction)
* [SVD intro](https://www.youtube.com/watch?v=P5mlg91as1c)
* [SVD](https://machinelearningmastery.com/singular-value-decomposition-for-machine-learning/)

**Phase 5: Model Validation & Selection**

**Topics Covered:**

* Cross-validation (K-Fold, StratifiedKFold)
* Bias-variance tradeoff
* Learning curves and validation curves

**Goals:**

* Evaluate models more reliably
* Detect underfitting/overfitting

**Resources:**

* [k-fold cross validation](https://www.youtube.com/watch?v=TIgfjmp-4BA)
* [bias, variance](https://www.youtube.com/watch?v=EuBBz3bI-aA)
* [cross validation theory](https://scikit-learn.org/stable/modules/cross_validation.html)
* [Bias-variance trade-off](https://towardsdatascience.com/machine-learning-bias-variance-tradeoff-and-regularization-94846f945131/)
* [Learning curves](https://scikit-learn.org/stable/auto_examples/model_selection/plot_learning_curve.html)