**Curriculum Plan (ODD SEM 2024): B.Sc. (H) Mathematics III Year (Semester V)**

**DSE-3(i): MATHEMATICAL DATA SCIENCE**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Dr. Tajender Kumar**  Assistant Professor  Department of Mathematics  Kalindi College (University of Delhi)  Delhi- 110008  Mobile: +91 7417837644  **E- mail**: [tajenderkumar@kalindi.du.ac.in](mailto:tajenderkumar@kalindi.du.ac.in) | |  | **Marks Distribution** | **Theory** | 90 Marks | |
| **Practical** | 40 Marks | |
| **Internal Assessment** | Assignment 30 Marks | |
|  | |
|  | |
| **Classes Assigned** | **Lectures** | 3 per week (Theory) | |
| **Practical** | 2 per week | |
| **References** |  | 1. Mertz, David. (2021). Cleaning Data for Effective Data Science, Packt Publishing.  2. Ozdemir, Sinan. (2016). Principles of Data Science, Packt Publishing.  3. Phillips, Jeff M. (2021). Mathematical Foundations for Data Analysis, Springer.  (https://mathfordata.github.io/). | | | | |
|  | **Week** | **Topics** | | | |  |
|  | **Beginning/1st week with 3 days**  **(**01-03,05-10AUG) | Types of Data: nominal, ordinal, interval, and ratio; Steps involved in data science case-study: question, procurement, exploration, modeling, and presentation; | | | |  |
|  | **2nd week (**12-17 AUG) | Types of Data: nominal, ordinal, interval, and ratio; Steps involved in data science case-study: question, procurement, exploration, modeling, and presentation; | | | |  |
|  | **3rd week (**19-24 AUG) | To deal with dirty and missing data, such as imputation, deletion, and data normalization. | | | |  |
|  | **4th week (**26-31 AUG) | To deal with dirty and missing data, such as imputation, deletion, and data normalization. | | | |  |
|  | **5th week (**02-07 SEP) | Use the real-world datasets (https://data.gov.in/) to demonstrate the following:  a) Data analysis and exploration, linear regression techniques such as simple, multiple  explanatory variables, cross-validation, and regularization. | | | |  |
|  | **6th week (**09-14 SEP) | Use the real-world datasets (https://data.gov.in/) to demonstrate the following:  a) Data analysis and exploration, linear regression techniques such as simple, multiple  explanatory variables, cross-validation, and regularization. | | | |  |
|  | **7th week (**16-21 SEP) | Use the real-world datasets (https://data.gov.in/) to demonstrate the following:  a) Data analysis and exploration, linear regression techniques such as simple, multiple  explanatory variables, cross-validation, and regularization. | | | |  |
|  | **8th week (**23-28 SEP) | Dimensionality reduction techniques such as principal component analysis, singular  value decomposition (SVD), and multidimensional scaling. | | | |  |
|  | **9th week (**30 SEP-05 OCT) | Dimensionality reduction techniques such as principal component analysis, singular  value decomposition (SVD), and multidimensional scaling. | | | |  |
|  | **10th week**. (07-12 0CT) | Problem of dimensionality, Principal component analysis, Singular value decomposition (SVD), | | | |  |
|  | **11th week (**14-19 0CT) | Clustering algorithms such as k-means, hierarchical, and density-based clustering and  evaluate the quality of the clustering results. | | | |  |
|  | **12th week (**21-26 OCT) | Clustering algorithms such as k-means, hierarchical, and density-based clustering and  evaluate the quality of the clustering results. | | | |  |
|  | **13th week (**04-09 NOV) | Classification methods such as linear classifiers, support vector machines (SVM), and  k-nearest neighbours (k-NN). [3]: Chapter 8. | | | |  |
|  | **14th week (**11-16 NOV) | Classification methods such as linear classifiers, support vector machines (SVM), and  k-nearest neighbours (k-NN). | | | |  |
|  | **15th week** (18-23 NOV) | Revision | | | |  |
|  | **16th week only with 2 Days** (25-27 NOV) | Revision | | | |  |
| Dispersal of classes, preparation leave and practical examination begin- 28 November, 2024. | | | | | | |