**Week – 1**

**Supervised Learning:** In supervised learning, the algorithm is trained on a labeled dataset, where the input data is paired with corresponding output labels. The goal is to predict new unseen data by mapping from inputs to outputs, making predictions or classifications.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| State-County FIPS | Total Active Borrowers | Very Low Income | Low Income | Married | American Indian/Alaskan Native | Asian | Black/African American | Hawaiin Pacific Islander | White | Hispanic | Average Number of Children | Average Household Size |
| 1001 | 44 | 12 | 32 | 15 | 0 | 0 | 21 | 0 | 19 | 0 | 1.1 | 2.6 |
| 1003 | 368 | 157 | 210 | 82 | 3 | 0 | 146 | 1 | 202 | 8 | 1.1 | 2.2 |
| 1005 | 124 | 34 | 91 | 23 | 0 | 0 | 108 | 0 | 12 | 0 | 1 | 2 |
| 1007 | 20 | 6 | 14 | 7 | 0 | 0 | 11 | 0 | 9 | 0 | 0.9 | 2 |

Here are a few supervised Learning Problems from the dataset of week-1:

1. **Regression - Predicting Total Active Borrowers:**
   * Target Variable: Total Active Borrowers
   * Features: Very Low Income, Low Income, Married, American Indian/Alaskan Native, Asian, Black/African American, Hawaiin Pacific Islander, White, Hispanic, Average Number of Children, Average Household Size.
2. **Regression - Predicting Average Number of Children:**
   * Target Variable: Average Number of Children
   * Features: Total Active Borrowers, Very Low Income, Low Income, Married, American Indian/Alaskan Native, Asian, Black/African American, Hawaiin Pacific Islander, White, Hispanic, Average Household Size.
3. **Classification - Predicting High vs. Low Household Size:**
   * Target Variable: High vs. Low Household Size (binary classification)
   * Features: Very Low Income, Low Income, Married, American Indian/Alaskan Native, Asian, Black/African American, Hawaiin Pacific Islander, White, Hispanic, Total Active Borrowers, Average Number of Children.
4. **Regression - Predicting Average Household Size:**
   * Target Variable: Average Household Size
   * Features: Total Active Borrowers, Very Low Income, Low Income, Married, American Indian/Alaskan Native, Asian, Black/African American, Hawaiin Pacific Islander, White, Hispanic, Average Number of Children.

**Unsupervised Learning:** In Unsupervised learning, the algorithm is trained on a unlabeled dataset, that searches the data for structures, relationships, or patterns without explicit instructions on what to look for. Analyzing the data's underlying structure or distribution often becomes the aim for this kind of dataset.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1013 | 179 | 81 | 97 | 35 | 1 | 0 | 141 | 0 | 26 | 0 | 0.9 | 1.9 |
| 1015 | 185 | 67 | 117 | 61 | 1 | 0 | 51 | 0 | 130 | 1 | 1.2 | 2.4 |
| 1017 | 109 | 43 | 64 | 18 | 0 | 0 | 85 | 0 | 19 | 0 | 0.9 | 2.2 |
| 1019 | 53 | 15 | 38 | 22 | 0 | 0 | 10 | 0 | 40 | 1 | 0.6 | 2.1 |

Here are a few Unsupervised Learning Problems from the dataset of week-1:

1. **Clustering - Grouping Counties Based on Demographics:**
   * Cluster counties based on demographic features such as Very Low Income, Low Income, Married, American Indian/Alaskan Native, Asian, Black/African American, Hawaiin Pacific Islander, White, Hispanic.
2. **Dimensionality Reduction - Reducing Features for Borrower Prediction:**
   * Use dimensionality reduction techniques to identify the most relevant features for predicting Total Active Borrowers.
3. **Association Rule Learning - Identifying Patterns in Demographics:**
   * Discover associations or patterns between demographic features (e.g., relationship between race and income levels).