**Machine Learning**

**IT4060**

A picture containing text, clipart, vector graphics

Description automatically generated

ASSIGNMENT 2

Predict H1N1 and Seasonal Flu Vaccines

|  |  |
| --- | --- |
| **STUDENT ID** | **NAME** |
| IT19125244 | Rogin Harshini N.V. |
| IT19074450 | Manilka G.S. |
| IT19364216 | P.J. Samuel |
| IT19038292 | Wickramasinghe A.Y.S.W. |

Table of Contents

[Problem Definition 2](#_Toc104455472)

[Dataset Description 3](#_Toc104455473)

[Methodology 4](#_Toc104455474)

[Results and Discussion 5](#_Toc104455475)

[Individual Contribution 6](#_Toc104455476)

[References 7](#_Toc104455477)

[Appendix 8](#_Toc104455478)

# Problem Definition

# Dataset Description

The dataset used is from the National H1N1 Flu Survey.

● Training features

● Training labels

● Testing features

|  |  |
| --- | --- |
| **Feature Name** | **Meaning** |
| h1n1\_concern | Level of concern about the H1N1 flu.   * 0 = Not at all concerned * 1 = Not very concerned * 2 = Somewhat concerned * 3 = Very concerned. |
| h1n1\_knowledge | Level of knowledge about H1N1 flu.   * 0 = No knowledge * 1 = A little knowledge * 2 = A lot of knowledge |
| behavioral\_antiviral\_meds | Has taken antiviral medications. |
| behavioral\_avoidance | Has avoided close contact with others with flulike symptoms |
| behavioral\_face\_mask | Has bought a face mask. |
| behavioral\_wash\_hands | Has frequently washed hands or used hand sanitizer. |
| behavioral\_large\_gatherings | Has reduced time at large gatherings. |
| behavioral\_outside\_home | Has reduced contact with people outside of their own household. |
| behavioral\_touch\_face | Has avoided touching eyes, nose, or mouth. |
| doctor\_recc\_h1n1 | H1N1 flu vaccine was recommended by a doctor. |
| doctor\_recc\_seasonal | Seasonal flu vaccine was recommended by the doctor. |
| chronic\_med\_condition | Has any of the following chronic medical conditions: asthma or another lung condition, diabetes, a heart condition, a kidney condition, sickle cell anemia or other anemia, a neurological or neuromuscular condition, a liver condition, or a weakened immune system caused by a chronic illness or by medicines taken for a chronic illness. |
| child\_under\_6\_months | Has regular close contact with a child under the age of six months |
| health\_worker | Is a healthcare worker. |
| health\_insurance | Has health insurance. |
| opinion\_h1n1\_vacc\_effectiv e | Respondent's opinion about H1N1 vaccine  Effectiveness.   * 1 = Not at all effective * 2 = Not very effective * 3 = Don't know * 4 = Somewhat effective * 5 = Very effective |
| opinion\_h1n1\_risk | Respondent's opinion about risk of getting sick with  H1N1 flu without a vaccine.   * 1 = Very Low * 2 = Somewhat low * 3 = Don't know * 4 = Somewhat high * 5 = Very high |
| opinion\_h1n1\_sick\_from\_v acc | Respondent's worry of getting sick from taking the  H1N1 vaccine.   * 1 = Not at all worried * 2 = Not very worried * 3 = Don't know * 4 = Somewhat worried * 5 = Very worried |
| opinion\_seas\_vacc\_effectiv e | Respondent's opinion about seasonal flu vaccine  effectiveness.   * 1 = Not at all effective * 2 = Not very effective * 3 = Don't know * 4 = Somewhat effective * 5 = Very effective |
| opinion\_seas\_risk | Respondent's opinion about risk of getting sick with  seasonal flu without a vaccine.   * 1 = Very Low * 2 = Somewhat low * 3 = Don't know * 4 = Somewhat high * 5 = Very high. |
| opinion\_seas\_sick\_from\_va cc | Respondent's worry of getting sick from taking the  seasonal flu vaccine.   * 1 = Not at all worried * 2 = Not very worried * 3 = Don't know * 4 = Somewhat worried * 5 = Very worried. |
| age\_group | Age group of respondents. |
| education | Self-reported education level. |
| race | Race of respondent. |
| sex | Sex of respondent. |
| income\_poverty | Household annual income of respondent with respect to 2008 Census poverty thresholds. |
| marital\_status | Marital status of respondent. |
| rent\_or\_own | Housing situation of respondents. |
| employment\_status | Employment status of respondent. |
| hhs\_geo\_region | Respondent's residence using a 10-region geographic classification defined by the U.S. Dept. of Health and Human Services. Values are represented as short random character strings. |
| census\_msa | Respondent's residence within metropolitan statistical areas (MSA) as defined by the U.S. Census. |
| household\_adults | Number of other adults in the household, topcoded to 3. |
| household\_children | Number of children in the household, top-coded to 3. |
| employment\_industry | Type of industry respondent is employed in. Values are represented as short random character strings. |
| employment\_occupation | Type of occupation of respondent. Values are represented as short random character strings. |

# Methodology

# Results and Discussion

# Individual Contribution

# References

# Appendix