

PES UNIVERSITY
Data Analytics- EC campus

Literature Survey Report

1.Project Title	Prediction of H1N1 Vaccine Uptake	
2.Team Name	Team Enigma	
3.Team Members	SRN1: PES2UG20CS415	Name1: Alheena Mujeeb
	SRN2: PES2UG20CS419	Name2: Apoorva Anand Kulkarni
4.Dataset used	National 2009 H1N1 Flu Survey (NHFS)	
5.Link for the Dataset	https://www.drivendata.org/competitions/66/flu-shot-learning/page/210	
6. Github link:	https://github.com/Alheena/H1N1-vaccine-uptake-prediction	

7.Problem Statement:

The dataset contains 33 attributes pertaining to social, economic, and demographic background, opinions on risks of illness and vaccine effectiveness etc. The goal of the project is to predict how likely individuals are to receive their H1N1 vaccines. A better understanding of how these characteristics are related to vaccination patterns can be useful for public health efforts

- **h1n1_concern** - Level of concern about the H1N1 flu.
- **h1n1_knowledge** - Level of knowledge about H1N1 flu.
- **behavioral_antiviral_meds** - Has taken antiviral medications. (binary)
- **behavioral_avoidance** - Has avoided close contact with others with flu-like symptoms. (binary)
- **behavioral_face_mask** - Has bought a face mask. (binary)
- **behavioral_wash_hands** - Has frequently washed hands or used hand sanitizer.
- **behavioral_large_gatherings** - Has reduced time at large gatherings. (binary)
- **behavioral_outside_home** - Has reduced contact with people outside of the household.
- **behavioral_touch_face** - Has avoided touching eyes, nose, or mouth. (binary)
- **doctor_recc_h1n1** - H1N1 flu vaccine was recommended by the doctor. (binary)
- **chronic_med_condition** - Has any of the following chronic medical conditions: asthma or an other 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. (binary)
- **child_under_6_months** - Has regular close contact with a child under the age of six months.
- **health_worker** - Is a healthcare worker. (binary)
- **health_insurance** - Has health insurance. (binary)
- **opinion_h1n1_vacc_effective** - Respondent's opinion about H1N1 vaccine effectiveness.
- **opinion_h1n1_risk** - Respondent's opinion about risk of getting sick with H1N1 flu without

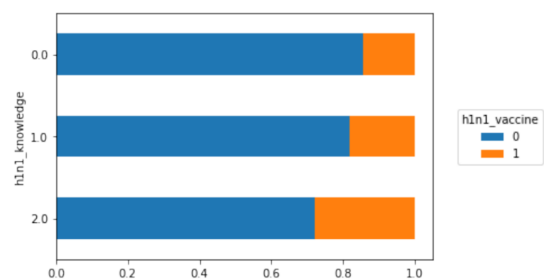
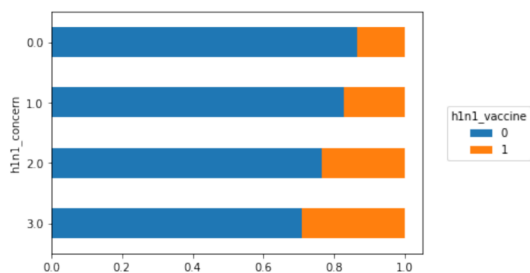
vaccine.

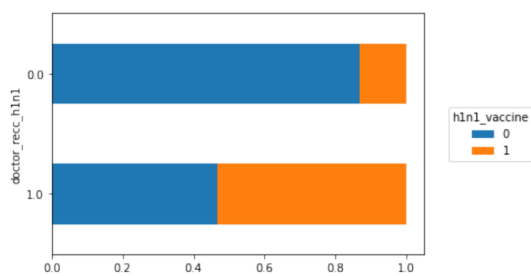
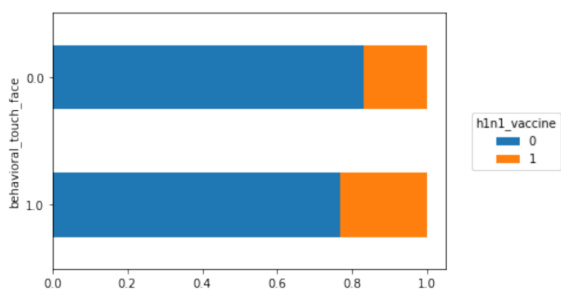
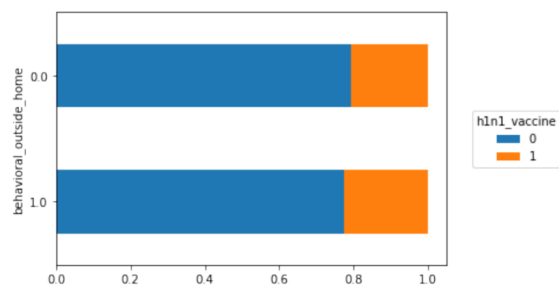
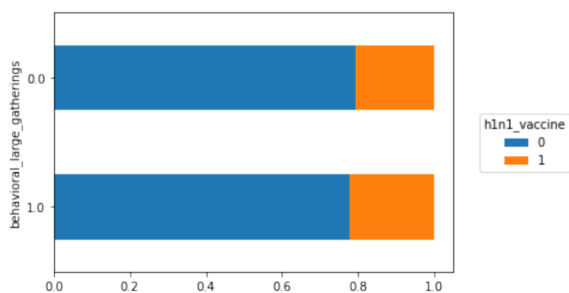
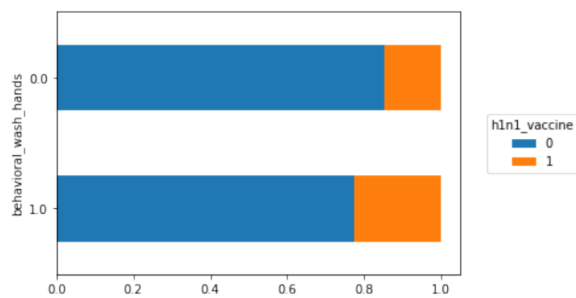
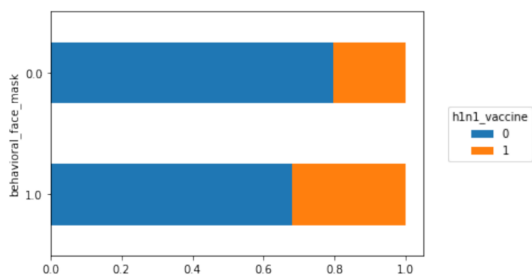
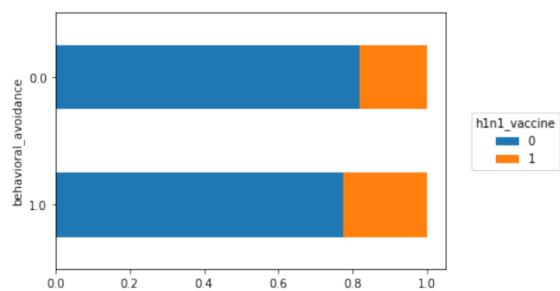
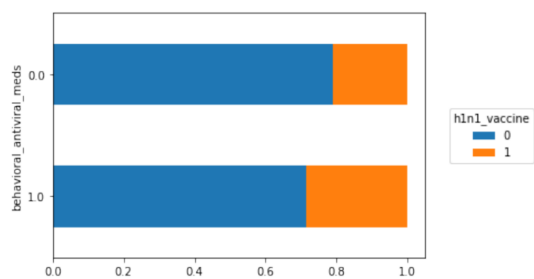
- **opinion_h1n1_sick_from_vacc** - Respondent's worry of getting sick from taking H1N1 vaccine.
- **age_group** - Age group of respondent.
- **education** - Self-reported education level.
- **race** - Race of respondent.
- **sex** - Sex of respondent.
- **income_poverty** - Household annual income of respondents with respect to 2008 Census poverty thresholds.
- **marital_status** - Marital status of respondent.
- **rent_or_own** - Housing situation of respondent.
- **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 household, top-coded to 3.
- **household_children** - Number of children in 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.

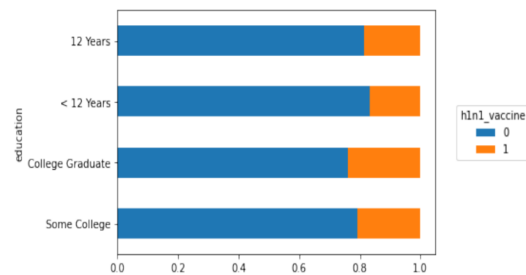
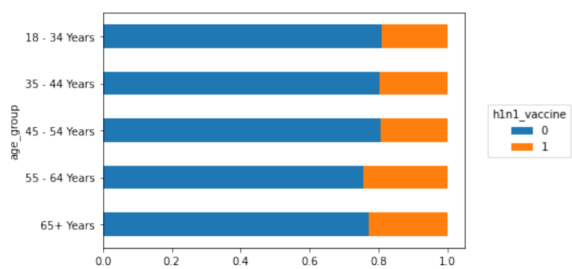
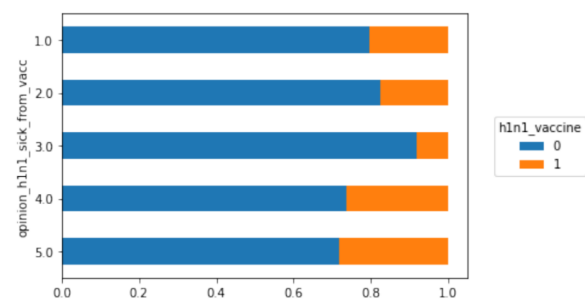
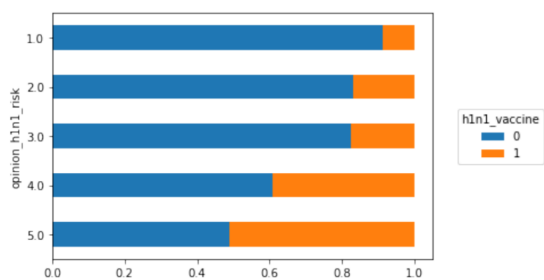
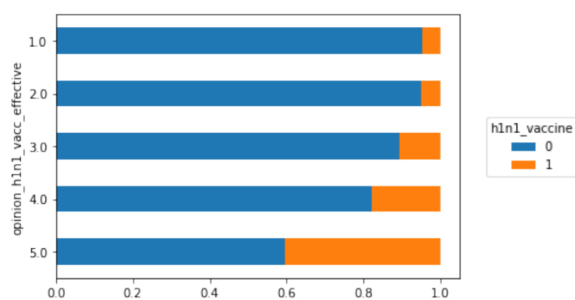
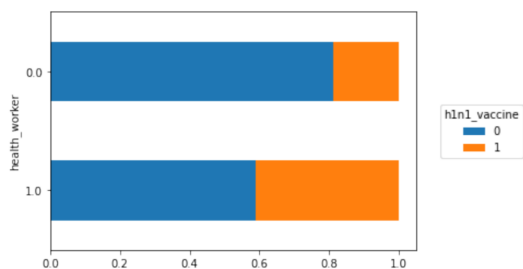
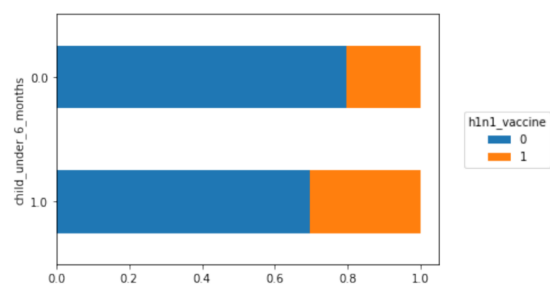
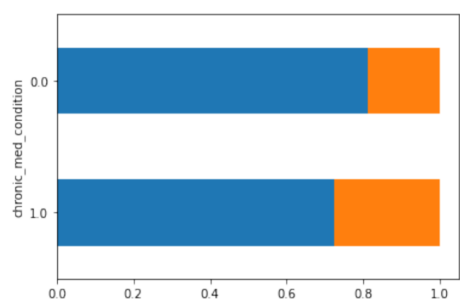
8.EDA and Visualization

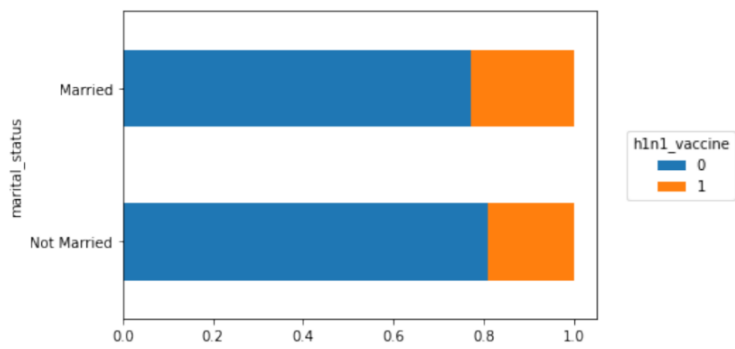
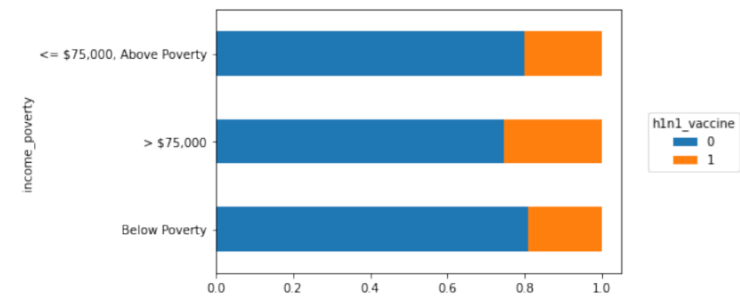
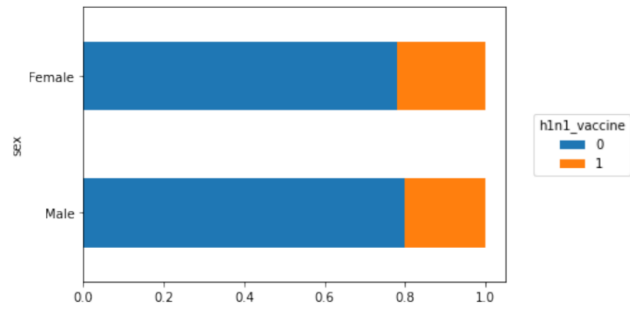
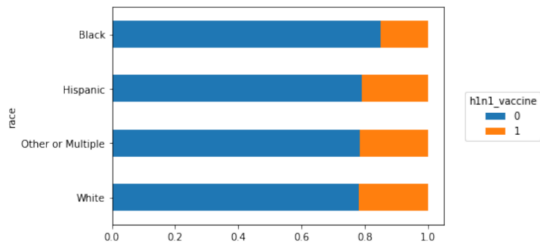
No of rows: 26707

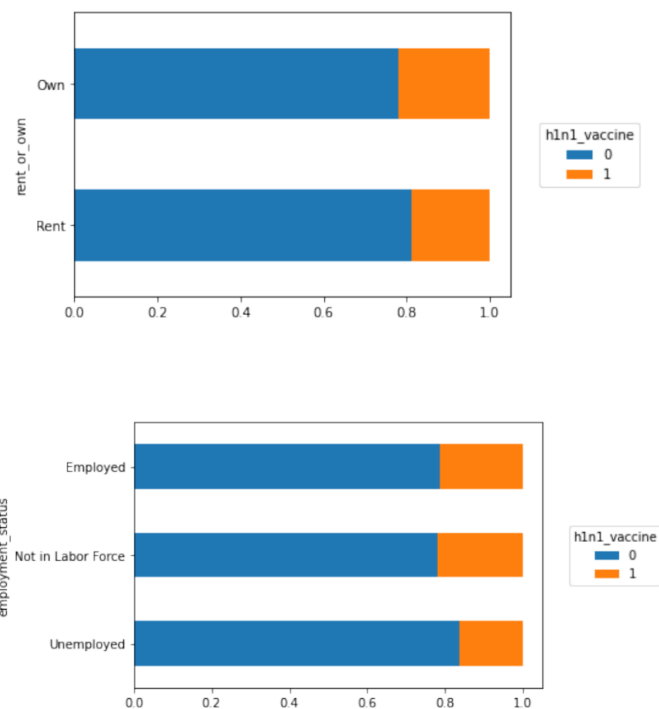
No of columns: 33











9. Summarize the Literature survey

1. This paper aims to determine the association between sociodemographic factors and vaccine uptake across counties in the United States. Data has been sourced from several online databases such as the US Centers for Disease Control and Prevention and the US Census Bureau COVID-19 Site. The authors have attempted to use the XGBoost model and have achieved an accuracy of 62%. Location, education, ethnicity, income, and household access to the internet have been identified as the most critical sociodemographic features. Hence, this paper reveals that sociodemographic characteristics are strong predictors of vaccine uptake rates across counties in the United States.
2. This paper attempts to identify predictors of willingness to accept a COVID-19 vaccine, and compares them to predictors of previous acceptance of an influenza vaccine among adults in the southeastern US. The dataset contains attributes related to health and healthcare utilization, COVID-19 testing, results, perceived risk of infection, COVID-19 vaccine willingness, and influenza vaccine acceptance. A series of univariate logistic regressions were employed to investigate and compare predictors of respondents' (1) willingness to accept a COVID-19 vaccine, and (2) previous acceptance of the influenza vaccine for the current flu season. This paper establishes a relationship between behaviors indicative of reducing the spread of COVID-19, education, income, recent healthcare utilization and vaccine uptake
3. The goal of this paper was to examine state-level features and policies that are most important

in achieving a threshold level vaccination rate to curve the effects of the COVID-19 pandemic. Chi-squared Automatic Interaction Detection, a Decision Tree algorithm (CHAID) was used to do the same. K-fold cross validation technique was utilized to determine algorithms' generalization error. 88% accuracy and a sensitivity score of 92.5% was achieved. Workplace travels, political affiliation of the governor, and the vaccine mandate in schools were the top three features to predict Covid 19 vaccine uptake.

4. In this paper "Seasonal and 2009 H1N1 influenza vaccine uptake, predictors of vaccination and self-reported barriers to vaccination among secondary school teachers and staff" it attempts to investigate factors associated with seasonal and 2009 H1N1 influenza vaccine uptake among middle- and high-school teachers in a rural, low-income setting during the 2009–2010 influenza season. They used chi-square test and attempted to correlate seasonal and 2009 vaccine uptake. Six psychosocial variables were assessed to determine participants' attitudes and beliefs regarding seasonal and H1N1 influenza vaccinations. Lastly by using the data gathered during the survey conducted, they concluded that teachers have a heightened risk of both contracting and spreading influenza to others, which can place their students and families at risk.
5. In this paper "H1N1 Influenza Vaccine Compliance among Hospital- and Non-Hospital-Based Healthcare Personnel". The main aim was to identify factors influencing non-hospital HCP H1N1 influenza vaccine compliance. In logistic regression stratified by hospital setting versus non-hospital setting, determinants of H1N1 vaccination among non-hospital-based HCP included extent to which H1N1 vaccination was mandated or encouraged, perceived importance of vaccination, access to no-cost vaccine provided on-site, no fear of vaccine side effects, and trust in public health officials when they say that the influenza vaccine is safe. Finally we concluded that Non-hospital-based HCP versus hospital-based HCP reasons for H1N1 vaccine uptake differed. Targeted interventions are needed to increase compliance with pandemic-related vaccines.
6. In this paper "Models to predict H1N1 outbreaks" the aim was to) To identify various mathematical models used in prediction of H1N1 outbreaks till date. 2) To compare the usefulness of these models in providing magnitude of H1N1 outbreaks. The data was collected from past influenza pandemics and 2009 influenza A (H1N1). Statistical methods such as Significant t-test was used Stochastic (probabilistic) and Deterministic (compartmental models, non probabilistic). Among the mathematical model 1) Basic reproduction number (R_0) for H1N1 2) SIER model for H1N1 3) SIS model for H1NI, 4) SER model for H1N1 5) SIQR Model for H1N1 6) MSIR model for H1N1, we found that SIER model is better. This :Used combination of mathematical models.

10. What is the specific problem your team is going to solve?

Firstly, we will attempt to determine the most important attributes that impact H1N1 vaccine uptake. To do so, we have read through papers regarding both H1N1 and Covid-19 vaccine uptake. Then, using these attributes we will attempt to build a model that will predict whether a person will take the vaccine given a set of attributes.

11. References	
Paper-1	<p>Cheong Q, Au-yeung M, Quon S, Concepcion K, Kong J Predictive Modeling of Vaccination Uptake in US Counties: A Machine Learning–Based Approach J Med Internet Res 2021;23(11):e33231 URL: https://www.jmir.org/2021/11/e33231 DOI: 10.2196/33231</p>
Paper-2	<p>Burch, A.E., Lee, E., Shackelford, P. et al. Willingness to Vaccinate Against COVID-19: Predictors of Vaccine Uptake Among Adults in the US. J Primary Prevent 43, 83–93 (2022). https://doi.org/10.1007/s10935-021-00653-0</p>
Paper-3	<p>Syed Muhammad Ishraque Osman, Ahmed Sabit, Predictors of COVID-19 vaccination rate in USA: A machine learning approach, Machine Learning with Applications, 2022, 100408, ISSN 2666-8270, https://doi.org/10.1016/j.mlwa.2022.100408.</p>
Paper-4	<p>Seasonal and 2009 H1N1 influenza vaccine uptake, predictors of vaccination and self-reported barriers to vaccination among secondary school teachers and staff by Lisa M. Gargano, Julia E. Painter, Jessica M. Sales, Christopher Morfaw, LaDawna M. Jones, Dennis Murray, Gina M. Wingood, Ralph J. DiClemente & James M. Hughes</p>
Paper-5	<p>Peiris, JSM, Poon, LLM, Guan, Y. Emergence of a novel swine-origin influenza A virus (S-OIV) H1N1 virus in humans. J Clin Virol 2009;45(3):169–173. World Health Organization. WHO recommendations on pandemic (H1N1) 2009 vaccines: pandemic (H1N1)</p>
Paper-6	<p>International Journal of Community Medicine and Public Health,,Mutalik AV. Models to predict H1N1 outbreaks: a literature review. Int J Community Med Public Health 2017;</p>