Team 3
Christopher Davis, Matthew Ward, Caleb Kelson, Mary Pulley, Madhavi Nithianandam, Katie Craig
TA - Revati

The Flockbusters

February 2025

Project description

Bird flu is currently spreading in the US resulting in a large public health risk due to zoonotic potential along with an economic impact to the poultry industry. We will assess the severity level of each strain using genomic analysis. The question we will answer is: **Which strains of bird flu are the most dangerous?**

Secondarily, we want to investigate **what is the prevalence of bird flu compared to all other types of flu in the United States?** As of February 5, 2025, the H5NI outbreak has spread to 48 states and impacted poultry along with cows[1]. We would like to understand how this compares to flu outbreaks over time and put it into perspective.

Why Machine Learning?

Machine learning is uniquely qualified to solve this problem because of the large volume of data required to answer our questions, while not sacrificing accuracy.

We will use the following approaches:

- K-Means, Principal Component Analysis(PCA)
- Classification (Random Forest, CNNS for sequence classification)

Dataset(s) and algorithms to be used

As we selected data, we noted that public, government health information may be limited or changing due to the current political climate. Access to some data sources may be at risk due to external factors.

Primary Data Source:

Center for Disease Control (CDC) - https://www.cdc.gov/bird-flu/situation-summary/index.html

[1] Leake, Lindsey. "Bird Flu FAQ: Everything You Need to Know About the H5N1 Outbreak That's Spread to Dairy Cows in 9 States." *Yahoo News*, February 5, 2025. https://www.yahoo.com/news/bird-flu-faq-everything-know-010000167.html.

Secondary Data Sources:

Global Initiative on Sharing All Influenza Data (GISAID) - https://gisaid.org/

World Health Organization (WHO) - https://www.who.int/

Data.gov - https://catalog.data.gov/dataset/?q=H5N1+&sort=views_recent+desc&tags=avian-influenza&ext_location=&ext_bbox=&ext_prev_extent=

Rough Breakdown of Tasks with Owners

Day 1

- [x] ~~Create Github Project Repo - Mary~~
- [x] ~~Create Project Charter - Katie~~
- [x] ~~Identify Data Sets - Madhavi, Matthew, Mary, Christopher~~
- [x] ~~Investigate Data in a notebook - Katie~~
- [x] ~~Submit Project Charter for review by TA - Katie~~

Day 2

- [ ] Determine the notebooks needed for the project
- [ ] Assign out who will work on each part
- [ ] Set next milestones to hit for Day 3