We all have an interest in public safety and have created governmental agencies to act on our behalf – airline safety (FAA) --- food and drug safety (FDA) – highway safety (DOT/FHSA). Bridge safety (BTS/FHA). And yet most of us never think about bridge safety until there is a disaster such as the Florida bridge collapse in 2018 that killed six (<https://www.usatoday.com/story/news/2018/11/15/ntsb-miami-bridge-collapse-design-errors/2012020002/>) or the Silver Bridge collapse in Ohio (1967) that killed 46 (<https://en.wikipedia.org/wiki/Silver_Bridge>). The National Bridge Inventory maintained by the FHA lists 616,087 bridges. 7.63% are rated “structurally deficient.” (47052/616087) and need urgent repairs. Estimated cost to complete all needed bridge work is $171 Billion. <https://artbabridgereport.org/>

Americans cross these deficient bridges 178 million times a day. (<https://artbabridgereport.org/reports/2019-ARTBA-Bridge-Report.pdf>).

The database I will be using is the:

FHA (Federal Highway Administration) National Bridge Inventory:

<https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm>

The **National Bridge Inventory** (**NBI**) is a database, compiled by the [Federal Highway Administration](https://en.wikipedia.org/wiki/Federal_Highway_Administration), with information on all [bridges](https://en.wikipedia.org/wiki/Bridge) and [tunnels](https://en.wikipedia.org/wiki/Tunnel) in the [United States](https://en.wikipedia.org/wiki/United_States) that have roads passing above or below. This is similar to the grade crossing identifier number database compiled by the [Federal Railroad Administration](https://en.wikipedia.org/wiki/Federal_Railroad_Administration) which identifies all railroad crossings. This bridge information includes the design of the bridge and the dimensions of the usable portion. The data is often used to analyze bridges and judge their conditions. The inventory is developed with the purpose of having a unified database for bridges to ensure the safety of the traveling public as required by the Federal-Aid Highway Act of 1968.[[1]](https://en.wikipedia.org/wiki/National_Bridge_Inventory#cite_note-NBIS-1) It includes identification information, bridge types and specifications, operational conditions, bridge data including geometric data and functional description, and inspection data. Any bridge more than 20 feet (6 meters) long used for vehicular traffic is included

Over 600,000 rows

Over 100 features

Most features have several sub-features.

The target Feature is feature 67 – “Structural Evaluation”

Every bridge is scored on a scale of zero to 10.

A score of 4 or less labels the bridge “Structurally Deficient”

Bridges are considered structurally deficient if they have been restricted to light vehicles, closed to traffic or require rehabilitation. Structurally deficient means there are elements of the bridge that need to be monitored and/or repaired. The fact that a bridge is "structurally deficient" does not imply that it is likely to collapse or that it is unsafe. It means the bridge must be monitored, inspected and maintained.

<https://www.virginiadot.org/info/resources/bridge_defs.pdf>

I have already spent many hours understanding the data – the meaning of each feature and its relationship to other features – and how they determine if a bridge is “Structurally Deficient.”

My objective is creating a model to predict which bridges are “structurally deficient.” using the raw data in a different way. I will run my model using the NBI database. I will score it against the bridges already identified as deficient.

My audience will be the general public but, in particular, individuals and entities responsible for bridge maintenance and safety.

I will, hopefully, present a method to prioritize bridge evaluations/inspections that is more granular, and cost effective, than the current practice.

MATH/ANALYSIS

After EDA – data cleaning, outlier detection, feature engineering -- I will subject the data to several ML models including Linear Regression, Random Forest and XGBoost. I will tune hyperparameters using Grid-Search -- ultimately recommending the model with the highest accuracy on the test data.