The glass classification dataset was initially employed in a paper, “Rule Induction in Forensic Science”, by Ian W. Evett and E.J. Spiehler in 1987. They recognized the usefulness to a forensic crime lab of classifying glass fragments from refractive index (RI) and chemical composition. This would, for example, allow the lab to ascertain whether samples gathered on a suspect’s clothing came from a window, potentially indicating he/she had broken it, or from another source, like a broken bottle.

For the experiment, Evett and Spiehler wished to see if a machine learning algorithm known as Bionic Evolutionary Algorithm Generating Logical Expressions, or BEAGLE, could correctly classify the glass, first as either window or non-window, and then into a second level of sub-categories. For this project, we are focused only on the window/non-window classification problem so will only discuss that portion of the original experiment.

To classify the glass, Evett and Spiehler trained the BEAGLE machine learning algorithm, which uses a series of “AND” statements to chain rules together based on the inputs. For example, a rule might look like: Fe (iron) <= Na (sodium) AND (K (potassium) > Fe\*650. They then tested the BEAGLE algorithm against two traditional statistical methods, k-nearest neighbor (k=3, in this case) and discriminant analysis.

To train the algorithm, Evett and Spiehler used a collection of 214 samples (the dataset we are using for this project) from the Home Office Forensic Science Laboratory in Birmingham, UK. After training, they then tested each of the three classification methods on this same dataset. The BEAGLE network performed best, with 8 incorrect classifications, versus 12 for each of the other two. The team then performed a blind trial on 10 further samples, with the BEAGLE algorithm and K-nearest neighbor returning 1 incorrect classification each, and discriminant analysis yielding 3. Evett and Spiehler concluded that the BEAGLE network showed promise as an inexpensive method, though further testing on more samples was needed.