Procedure for Categorizing Lithology Information

Once the lithologic descriptions are machine-readable, they must be sorted into a set number of categories that will be used in the interpretation of the AEM data, the development of the rock physics transform, and the construction of the HCM. As an example, for the work in Butte County, sediment texture was used, defining four categories: 1) “fine” for fine-dominated sediment packages such as clay, 2) “mixed” for mixed fine- and coarse sediment packages like silty sands or sandy clays, 3) “coarse” for coarse-dominated sediment packages such as sand or gravel, and 4) “others” for descriptions where it is unclear what the driller was referring to.

We highly recommend that those responsible for this portion of the workflow develop a machine learning based method to complete this step. N. Dewar developed a method using a machine learning method referred to as Long Short Term Memory (LSTM) recurrent neural nets (RNNs) to automatically sort the lithologic descriptions into the defined categories. Using a set of lithologic descriptions that are sorted as training data, the LSTM RNNs are trained and are then able to accurately and instantly sort new lithologic descriptions, provided there is information in the original set of training data that allows for the sorting of the new lithologic descriptions.