Machine Learning in GENESIS

The Boosting

The term 'Boosting' refers to a family of algorithms which converts weak learner to a strong learner.

Let's understand this definition in detail by solving a problem of exploration identification. For example, you have a surface block model containing geophysical, geomorphological, and geochemical data, and all blocks contain surface data. Some zones are drilled and show no interesting mineralization, these blocks are marked "0", other zones are drilled and have interesting mineralization, these blocks are marked "1", other blocks are not drilled and marked "-1" for unknown.

Any surface information makes it possible to deduce somewhat the presence or not of the interesting mineralization. On the other hand, knowing the exact importance of each piece of information (weak learner) to make a good prediction and how to optimally combine this information to make a good prediction (strong learner) on the unknown zones is a very difficult problem.

The boosting allows you to know the importance of each of this information and you can choose to invest optimally in geophysics, geomorphology or geochemistry. There is no need to invest in expensive analyzes for an insignificant information supplement.

More importantly, boosting makes it possible to make predictions on unknown zones and to test the quality of these predictions on known zones. Genesis offers a learning and testing tool allowing to carry out all the combinations of learning on all the known zones except one conserved for the test. This makes it possible to precise the error on the predictions.

The Phylogenetics

In biology, phylogenetics is the study of the evolutionary history and relationships among individuals or groups of organisms (e.g. species, or populations). These relationships are discovered through phylogenetic inference methods that evaluate observed heritable traits, such as DNA sequences or morphology under a model of evolution of these traits. The result of these analyses is a phylogeny (also known as a phylogenetic tree) – a diagrammatic hypothesis about the history of the evolutionary relationships of a group of organisms.

As in biology, geomorphology and geochemistry are often evolutive processes which different current states derive from an identical previous state. The algorithms of phylogenetics take advantage of this knowledge to realize a tree of the evolution of these processes. This allows us to shed light on the complex mechanisms underlying geomorphology or current geochemistry.