
**CLASIFICACIÓN DE LITOLOGÍA Y ESTIMACIÓN DE VETAS EN TESTIGOS DE
SONDAJES DIAMANTINO USANDO MACHINE LEARNING Y VISIÓN
COMPUTACIONAL**

**SEBASTIÁN IGNACIO ARRIAGADA SILVA
INGENIERO CIVIL EN MINAS**

RESUMEN

La exploración geológica y la descripción de muestras de corteza terrestre son de suma importancia para la evaluación de yacimientos minerales. Por esta razón, existen múltiples esfuerzos por mejorar la metodología de descripción o de mapeo de rocas, siendo los algoritmos de machine learning basados en redes neuronales una potencial herramienta para estandarizar, automatizar y agilizar estos procesos. Este trabajo aborda la aplicación de dos tipos de modelos de machine learning, con el objetivo de clasificar la litología y estimar la presencia de vetas utilizando fotografías digitales de sondeos tipo diamantino. El primer tipo de modelo es generado a través del proceso Automated Machine Learning (AutoML), el cual adapta modelos de clasificación preexistentes con buenos rendimientos en bases de datos estandarizadas y los optimiza para ser entrenados en una nueva base de datos. La clasificación de imágenes fue utilizada para la inferencia de la litología, del estado geotécnico y de la presencia de vetas. El segundo tipo, correspondiente un modelo cGAN (Conditional Generative Adversarial Networks), está compuesto por dos redes neuronales enfrentadas en un juego de suma cero, en donde si uno gana el otro pierde, obteniendo destacados resultados en tareas como la transferencia de estilos en imágenes. Este modelo fue utilizado para la segmentación de mineralización en vetas. Para cuantificar el desempeño de los modelos se utilizaron 5 métricas derivadas de las correspondientes matrices de confusión. Con respecto a los resultados de la clasificación litológica, el mejor rendimiento se consiguió utilizando la etiqueta binaria correspondiente a "litología de interés" y "otro tipo", junto con entrenar un modelo diferente especializado en cada estado geotécnico presente (roca fracturada, molida, intacta y semi intacta). Se obtiene en conjunto un acierto del 79% de las veces y un 59% de posibilidades de que al reconocer la litología de interés, esta sea correctamente etiquetada. El

mayor rendimiento se obtiene dentro de la roca intacta con 78% y 79% respectivamente. Al analizar los datos clasificados erróneamente según la información obtenida por observación directa (verdad fundamental), se detectan casos que corresponden a potenciales zonas de transición entre la litología de interés y otras. Por lo tanto, si bien aún no se obtiene un nivel de asertividad necesario para automatizar por completo el proceso de mapeo, esta metodología presenta una potencial utilidad para realizar un perfil coherente de cada litología y detectar errores dentro de mapeos anteriores. Con respecto a la estimación de vetas, al segmentar las zonas de mineralización observable en las fotografías, es posible estimar su área con una exactitud del 97% y una precisión del 83% con respecto a los datos de prueba segmentados manualmente por un humano. La efectividad de estimación se ve reducida por factores como la variabilidad de tonalidades a lo largo de la roca y la presencia de zonas de alto brillo en la superficie. Aplicando ambas metodologías, es posible realizar inferencias dentro de nuevos sondajes a una rapidez por sobre 1 metro por segundo. Se recomienda utilizar esta herramienta como apoyo para agilizar el proceso de mapeo, pre clasificando y segmentando las zonas con altos niveles de certeza y señalar las zonas con mayor incerteza para la validación manual por parte del equipo de geología.

ABSTRACT

Geological exploration and description of earth's crust samples are of prime importance for the evaluation of mineral deposits. For this reason, there are multiple efforts to improve the methodology of description or rock mapping, being machine learning algorithms based on neural networks a potential tool to standardize, automate and streamline these processes. This work addresses the application of two types of machine learning models, with the objective of classifying lithology and estimating the presence of veins using digital photographs of diamond drill holes. The first type of model is generated through the Automated Machine Learning (AutoML) process, which adapts pre-existing classification models with good performance in standardized databases and optimizes them to be trained on a new database. The image classification was used for inference of lithology, geotechnical state and presence of veins. The second type, corresponding to a cGAN (Conditional Generative Adversarial Networks) model, is composed of two neural networks facing each other in a zero-sum game, where if one wins the other loses, obtaining outstanding results in tasks such as the transfer of styles in images. This model was used for the segmentation of mineralization in veins. To quantify the performance of the models, 5 metrics derived from the corresponding confusion matrices were used. With respect to the lithological classification results, the best performance was achieved by using the binary label corresponding to "lithology of interest" and "other type", together with training a different model specialized in each geotechnical state present (fractured, ground, intact and semi-intact rock). Overall a 79% success rate is obtained and a 59% chance that when recognizing the lithology of interest, it is correctly labeled. The highest performance is obtained within intact rock with 78% and 79% respectively. When analyzing the misclassified data according to the information obtained by direct observation (ground truth), cases are detected that correspond to potential transition zones between the lithology of interest and others. Therefore, although a level of assertiveness necessary to fully automate the mapping process is not yet obtained, this methodology presents a potential utility to make a coherent profile of

each lithology and to detect errors within previous mappings. With respect to vein estimation, by segmenting the zones of observable mineralization in the photographs, it is possible to estimate their area with an accuracy of 97% and an accuracy of 83% with respect to test data manually segmented by a human. The estimation effectiveness is reduced by factors such as the variability of shades along the rock and the presence of high brightness zones on the surface. By applying both methodologies, it is possible to make inferences within new drillholes at a speed of over 1 meter per second. It is recommended to use this tool as a support to speed up the mapping process, pre-classifying and segmenting the zones with high levels of certainty and pointing out the zones with higher uncertainty for manual validation by the geology team.