

Entrega 1: Modelo de clasificación de hongos venenosos

Profesor:

Raúl Ramos Pollán

Integrante:

Johan Daniel Zuluaga Gómez

Contacto:

johand.zuluaga@udea.edu.co

Universidad de Antioquia

Facultad de Ingeniería

2023

Contexto de la aplicación

En Colombia, donde el clima se caracteriza por condiciones húmedas y cálidas, los hongos prosperan en diversos ecosistemas, incluidos bosques tropicales, bosques de gran altitud y regiones tropicales. Los hongos desempeñan un papel vital en los ecosistemas del país, contribuyendo al ciclo de nutrientes y a la biodiversidad general de la región. Sin embargo, además de su importancia ecológica, los hongos tienen una considerable importancia cultural y económica.

Uno de los aspectos clave de este proyecto es abordar el valor nutricional de los hongos en el contexto de las comunidades que enfrentan problemas de alimentación. Muchas comunidades en Colombia dependen de la búsqueda de hongos como fuente de alimento, especialmente en áreas rurales donde el acceso a otros alimentos ricos en proteínas puede ser limitado. Los hongos pueden ser una fuente valiosa de nutrientes, incluidas proteínas, vitaminas y minerales, y potencialmente aliviar las deficiencias nutricionales en poblaciones vulnerables.

Sin embargo, buscar hongos silvestres conlleva riesgos importantes, ya que muchos hongos son tóxicos y pueden provocar una intoxicación grave o incluso la muerte si se consumen. Por lo tanto, la clasificación precisa de los hongos en categorías comestibles y venenosas es de suma importancia para el bienestar y la seguridad de la comunidad.

Objetivo de Machine Learning

Crear un modelo que prediga si un hongo es comestible o venenoso dada una imagen como entrada.

Dataset

Se usará el dataset de Kaggle [*edible and poisonous fungi*](#) que está dividido en 4 clases, hongos con y sin esporocarpos, que son las estructuras de los hongos que permiten la generación de esporas como su forma de reproducción sexual; cada una de estas clases se divide en comestible y venenosa. Se tiene la intención de usar las últimas dos clases mencionadas. El dataset ocupa un tamaño de 270.97 MB en disco, se compone de 3401 imágenes de diferentes tamaños y formatos de los cuales incluyen jpg, jpeg y png. La cantidad de imágenes que corresponden a la clase de hongos comestibles es de 1181, mientras que de la clase de hongos venenosos se tienen 2220 imágenes, por lo que se puede notar que hay un desbalance del doble.

Métricas de desempeño

Accuracy: la métrica principal para evaluar el rendimiento del modelo es la precisión, que mide la proporción de imágenes de hongos clasificadas correctamente.

Matriz de confusión: para comprender mejor los errores de clasificación y el comportamiento del modelo en diferentes clases.

Como métrica de negocio se podría calcular la cantidad de dinero que se puede perder en clasificar hongos comestibles en hongos venenosos, siempre prefiriendo que el modelo se equivoque en esta clasificación en lugar de la contraria dado a que podría llevar a resultados fatales, por eso también se recomienda que el modelo sea usado como apoyo a un experto que pueda identificar estos errores y evitar desgracias.

Referencias y resultados previos

- L. Picek, M. Šulc, J. Matas, J. Heilmann-Clausen, T. S. Jeppesen, and E. Lind, "Automatic Fungi Recognition: Deep Learning Meets Mycology," *Sensors*, vol. 22, no. 2, p. 633, Jan. 2022, doi: 10.3390/s22020633. [[Google Scholar](#)]
- M.-E. Nilsback and A. Zisserman, "Automated Flower Classification over a Large Number of Classes," in *Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, Washington, DC, USA, Dec. 16-19, 2008, pp. 722-729. [[Google Scholar](#)]
- M. W. Tahir, N. A. Zaidi, A. A. Rao, R. Blank, M. J. Vellekoop and W. Lang, "A Fungus Spores Dataset and a Convolutional Neural Network Based Approach for Fungus Detection," in *IEEE Transactions on NanoBioscience*, vol. 17, no. 3, pp. 281-290, July 2018, doi: 10.1109/TNB.2018.2839585. [[Google Scholar](#)]