

How to classify SECOP contracts according to the contract description?

Juan C. Gallego, Daniel F. Grijalba, Julian D. Hernández, Tomas Lemus, Alexander Ortega, Yovanni Rojas, Flor M. Vela.

Highlights

- Model based on a group of neural networks that interact sequentially, each one custom-designed, to generate the final classification.
- Our solution contributes to making public hiring processes more effective and efficient, reducing the user's search times of UNSPSC codes, and reducing their probability of error.
- Simple to use: write the object of your contract. Click on predict, and voila.

Problem

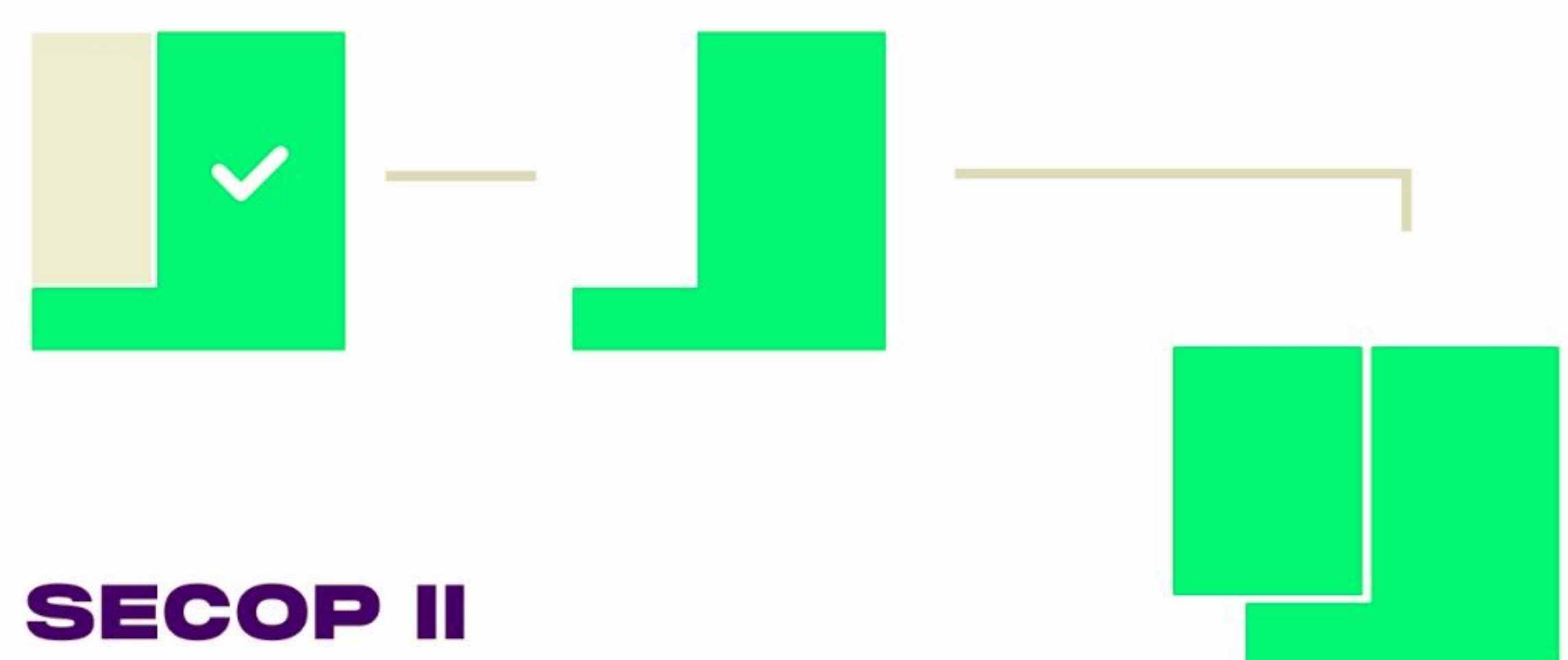
When a user is going to add a new contract, they must enter the description of the goods or services needed and classify it in the correct group according to its description. Classifying codes can be confusing for the user and can lead to misclassification. The correct classification of contracts is crucial for an efficient public contraction system.

Data

Data consists the available information about 6 million historical nationwide state contracts, grouped in the following formats:

- SECOP Electronic contracts (I & II)
- SECOP PAA (I & II)

SECOP I



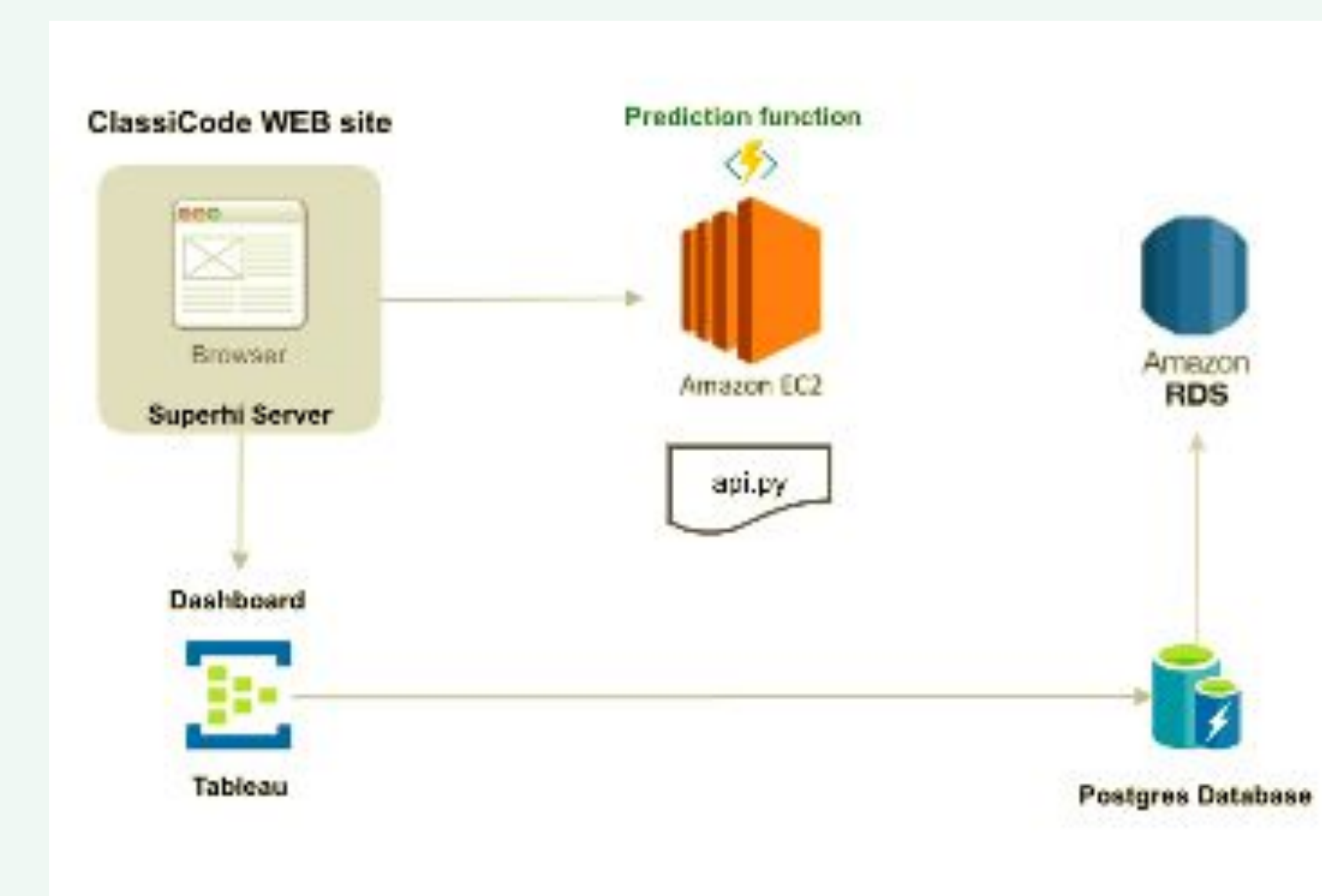
SECOP II



Final Dataset configuration based on SECOP (I & II)

ClassiCode

A Machine Learning Application build to assist the public contracting process. It is based on supervised neural networks and it's used to predict the United Nations Standard Products and Services Code - UNSPSC.

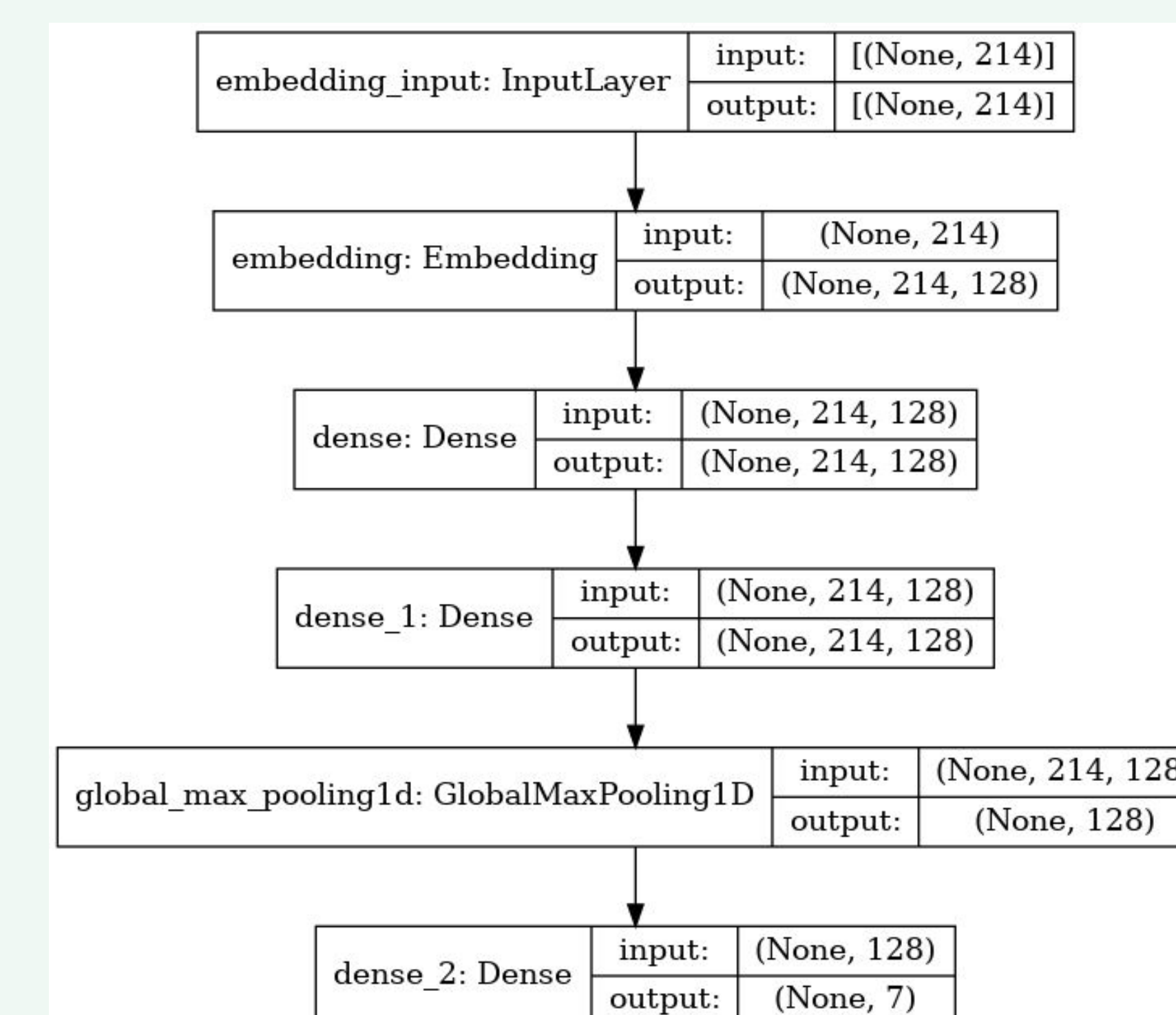
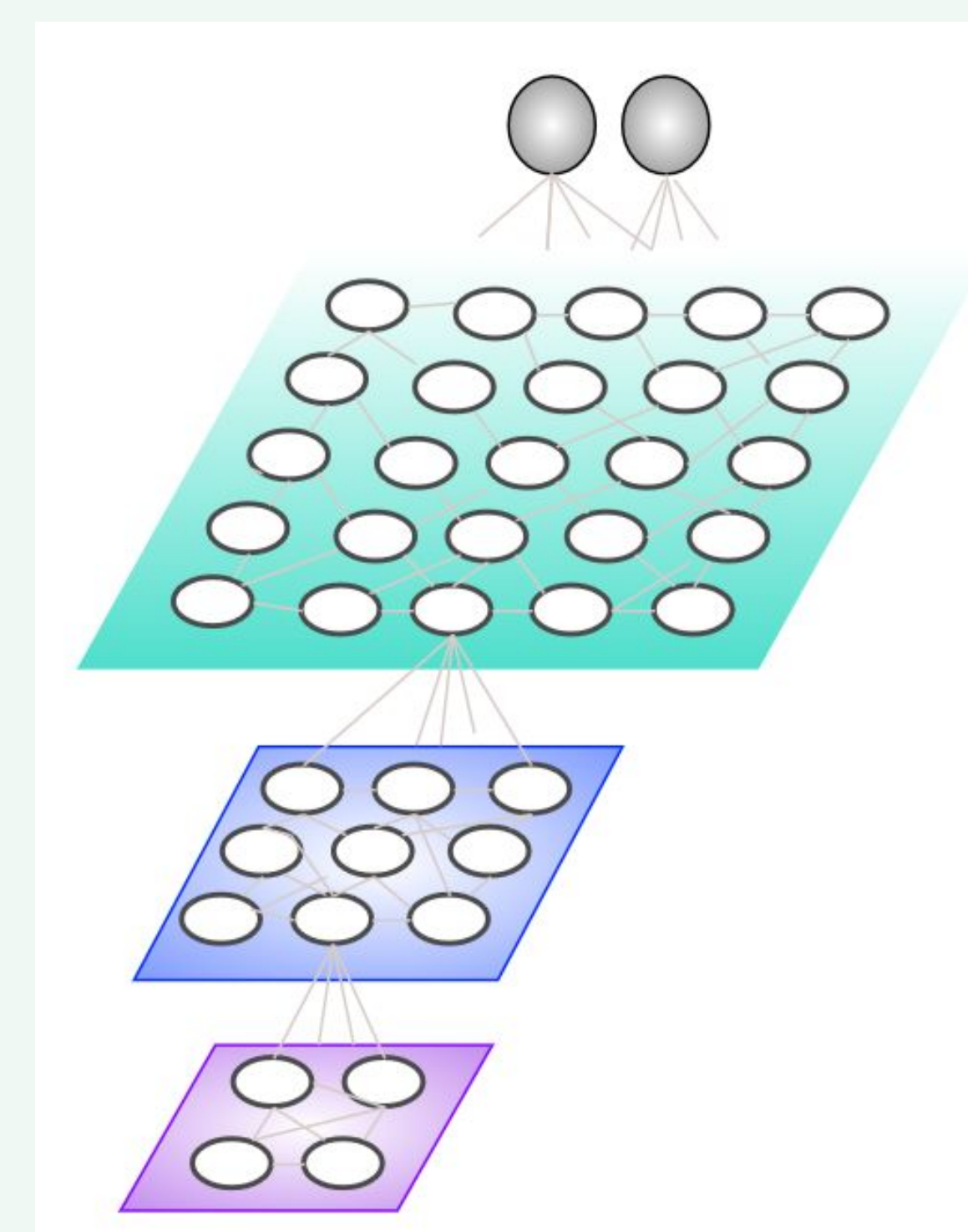


ClassiCode Architecture

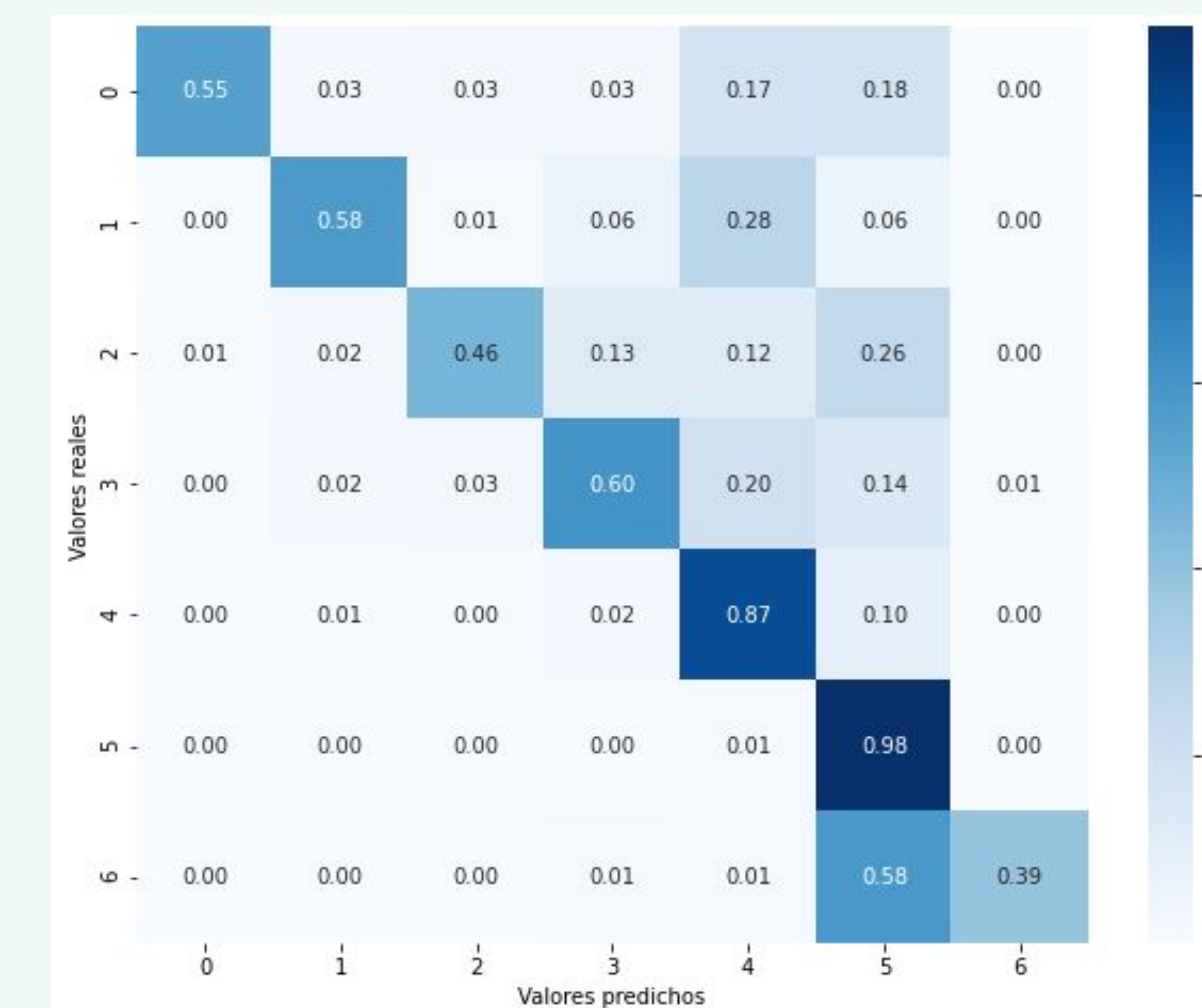
Hierarchical Neural Network Model

Offers flexibility and good performance to find relationships between selected characteristics. Also, it performs well on large volumes of data.

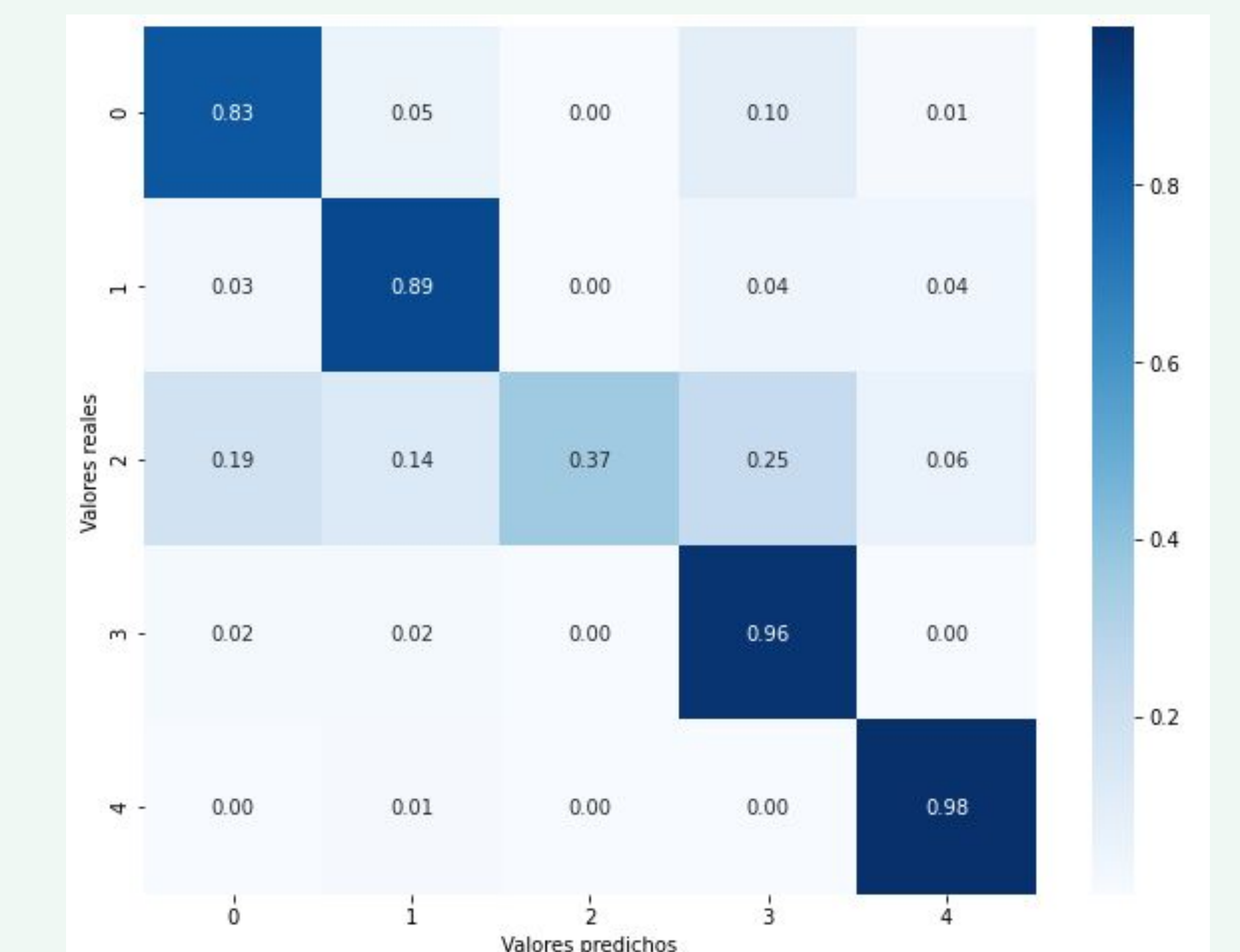
The model has an accuracy of 94%, an average precision of 63%, and an average recall of 76%.



Classi Code classifies UNSPSC segment codes and groups with an accuracy of 70% - 90%



Percentage confusion matrix with respect to the number of total records per group



Percentage confusion matrix with respect to the number of total records per segment of group B

A small example of how ClassiCode output is displayed.

ClassiCode

Escriba la descripción del contrato, haga clic en el botón. Obtendrá la predicción.

Contratación de desarrollo de página web para secretaria de agricultura.

Primer grupo sugerido
F: Servicios
Segmentos con mayor probabilidad
81: Servicios Basados en Ingeniería, Investigación y Tecnología
80: Servicios de Gestión, Servicios Profesionales de Empresa y Servicios Administrativos
82: Servicios Editoriales, de Diseño, de Artes Graficas y Bellas Artes

Predicir