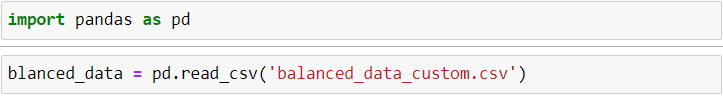
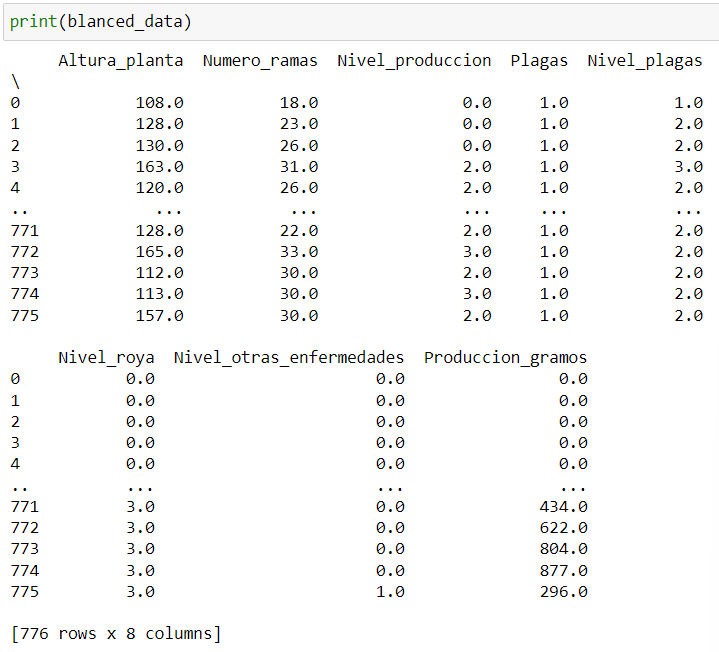
Para realizar la predicción de la roya mediante la utilización de características fenotípicas el autor de ese trabajo de titulación afirma que el modelo SVM (Support Vector Machine) es el que presenta un mejor rendimiento al momento de utilizarlo para investigaciones del ámbito agrícola. El desarrollo del modelo se muestra a continuación:

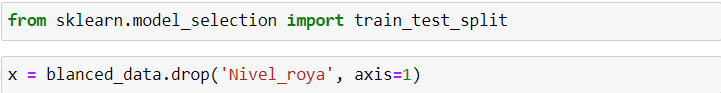
En primer lugar, se cargan los datos a utilizar, para ello se hace uso de la librería de Python denominada pandas, que en este caso es usada para leer los datos en el formato CSV (Valores separados por comas).



A continuación, se muestra el conjunto de datos cargados, donde existen 8 campos de acuerdo a cada una de las características dadas por el autor y 776 registros de cada uno.

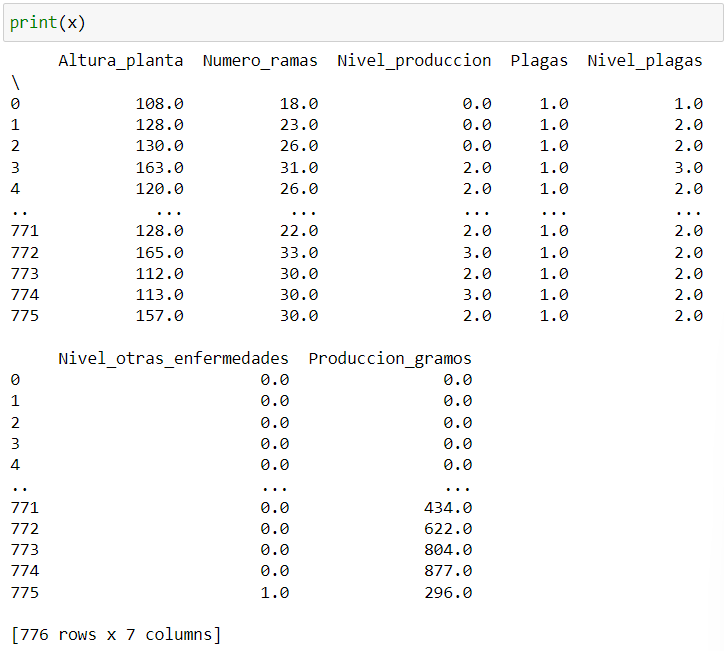


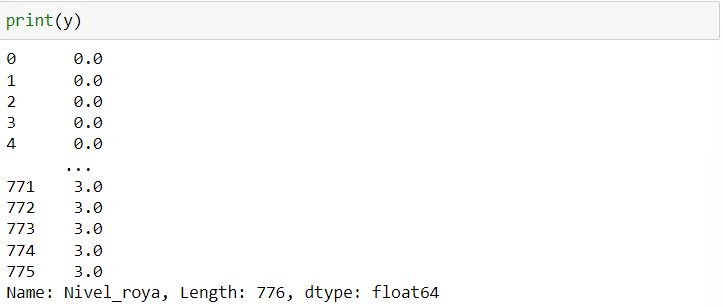
En segundo lugar, se separó la variable dependiente (‘Nivel\_roya’) de todo el conjunto de las otras variables, para utilizarlo en el modelo.





Siendo **x**, todas las variables independientes, que son las que el modelo va a solicitar para su ejecución y la **y** para realizar la comprobación de si el modelo funciona o no.

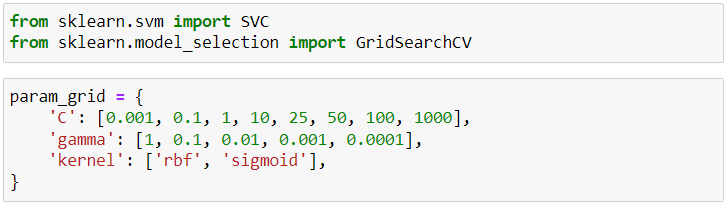




En tercer lugar, se separaron los datos para entrenamiento y para test, en dos subconjuntos donde el 80% corresponde para el primero y el 20% para el segundo, para ello se utilizó la siguiente función:



Antes de utilizar el modelo se realiza una optimización de hiper parámetros, utilizando para ello la función GridSearchCV la cual permite evaluar y seleccionar de forma sistemática los parámetros de un modelo y con ello obtener la mejor configuración para un modelo optimizado. Donde se definen los posibles valores de parámetros en un diccionario que luego va a ser usado en conjunto con el modelo.



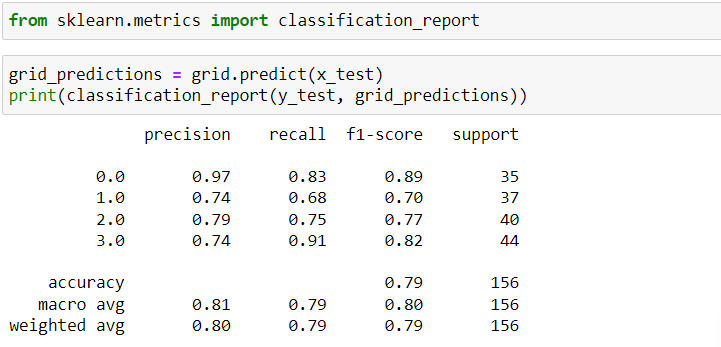
Para este modelo se utilizaron los parámetros:

* ‘C’ es un parámetro crítico de penalización que pueden tomar un rango de valores y tiene un efecto dramático en la forma de las regiones resultantes para cada clase.

Teniendo todos los parámetros listos se hace uso de la funcionalidad fit que permite realizar el entrenamiento de un modelo, en este caso el SVC en conjunto a sus parámetros a evaluar.



Una vez entrenado el modelo y obtenida la mejor configuración, el siguiente paso es probarlo y para ello se hace uso de los datos de test que corresponden al 20% que fueron separados del conjunto inicial. Para ello se hace uso de un reporte de clasificación, el cual permite visualizar el comportamiento del modelo al utilizar el testeo.



Con una precisión (accurary) del 79% de aciertos, siendo un porcentaje relativamente bueno al momento de utilizar este modelo de inteligencia computacional para predecir la roya en café robusta.