

**Instituto Tecnológico y de Estudios Superiores de Monterrey**

---

**06 de junio**

# **Machine learning with Python**

---

## **Profesores:**

Jose Gerardo Tamez

Martha Rebeca Canales

## **Grupo 301**

## **Integrantes del equipo:**

Andrea Corrales Romero A00828752

Juan Diego García Manrique A00829257

Rael Alejandro Barragán Reyes A01411256

Isabela Reséndez Sepúlveda A01194082

**06/06/2022**

**Monterrey, Nuevo León**

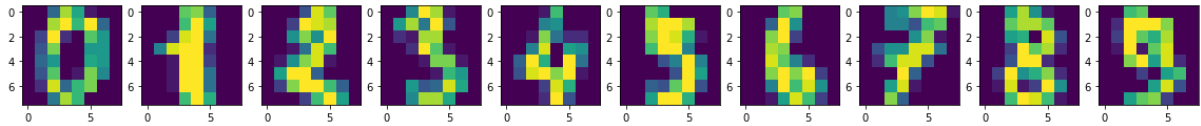
**SEMESTRE FEB - JUN 2022**

Con la clasificación de los dígitos por medio de machine learning se analizaron cuatro métodos diferentes, el método KNN, Naive Bayes, CART y el método LDA, a continuación se muestra un display de los números evaluados:

Labels' size: (1797,)

Features' size: (1797, 64)

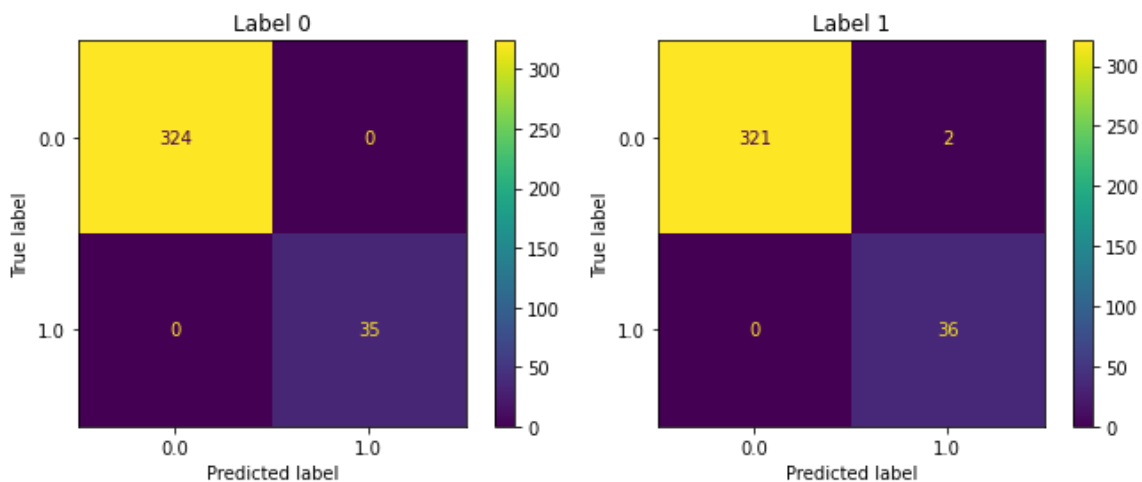
Images' size: (1797, 8, 8)



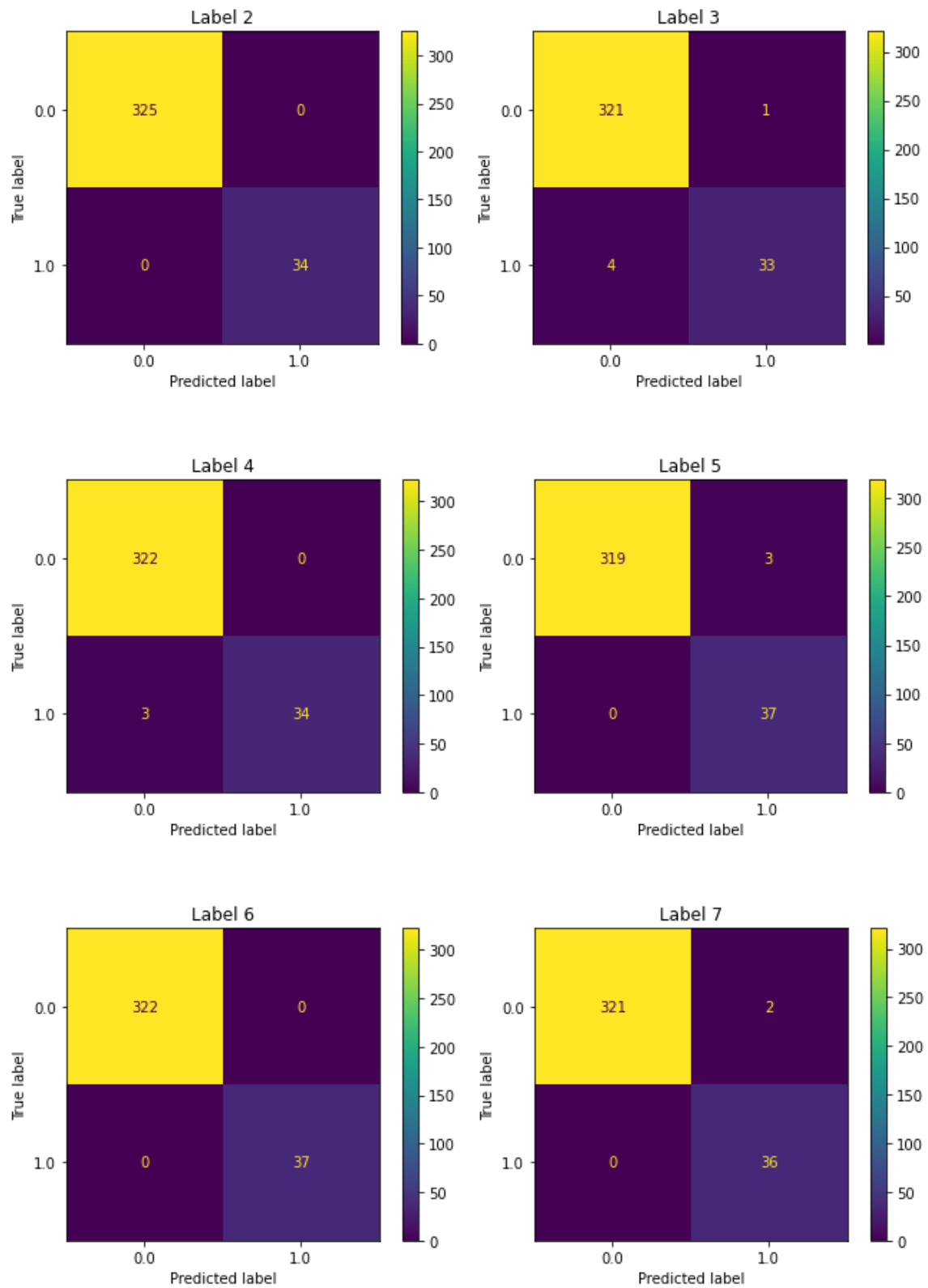
Después de obtener los números que serán evaluados con los métodos se comienza a hacer un entrenamiento al machine learning y a testear los sets. Una vez probados los ejemplos de entrenamiento se elige el método para hacer la prueba de exactitud de detección de los números.

## Resultados

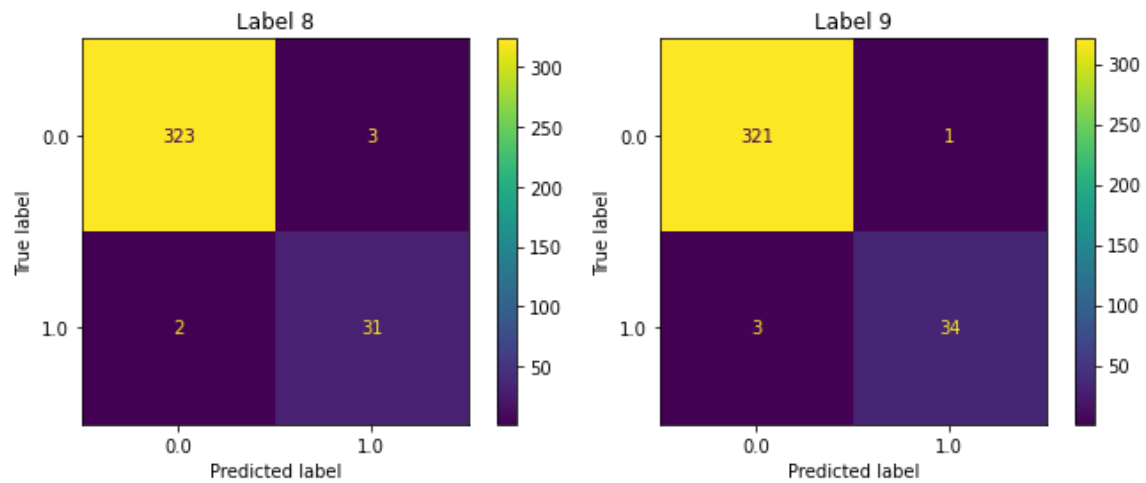
**KNN method:** The accuracy is: 96.66 %.



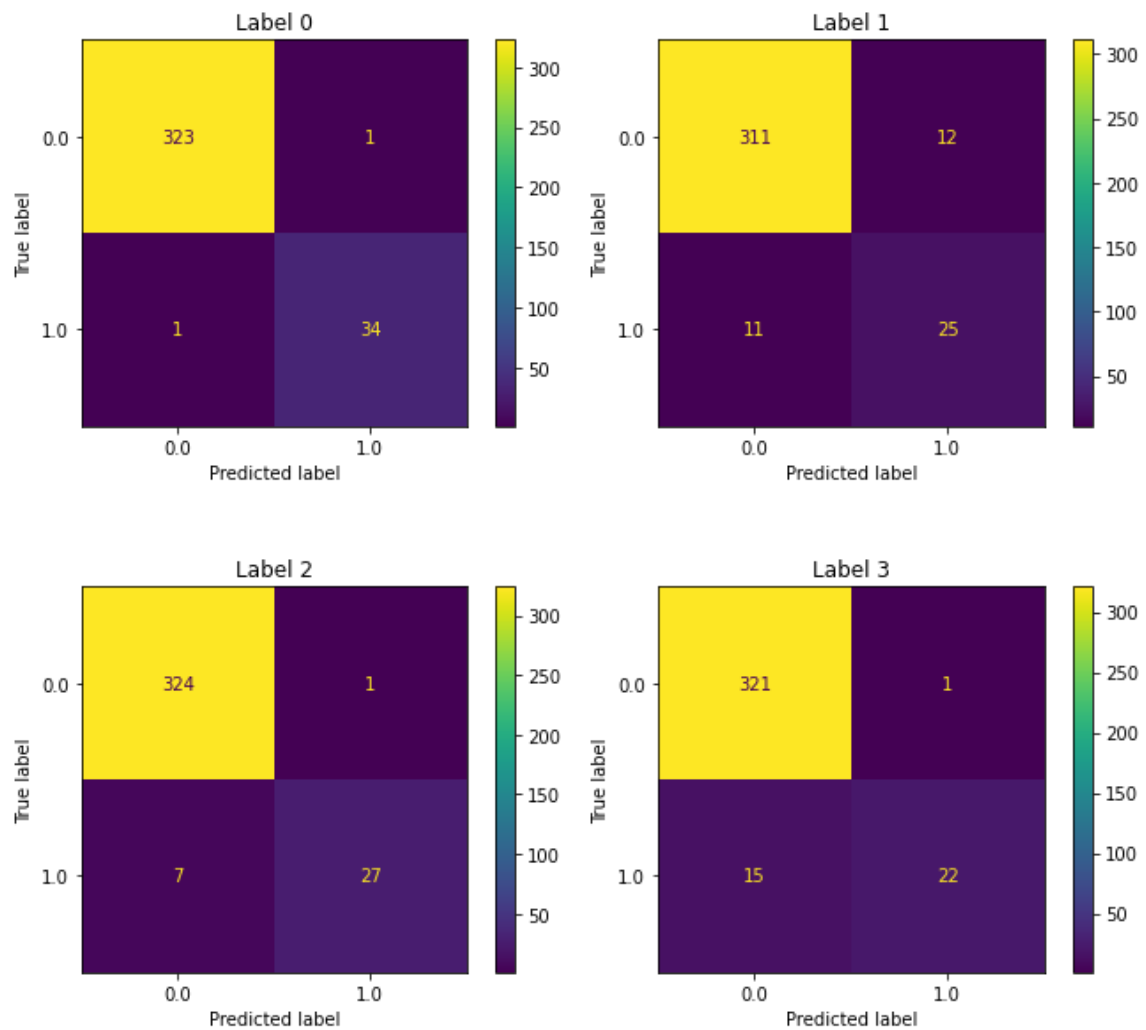
## Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)



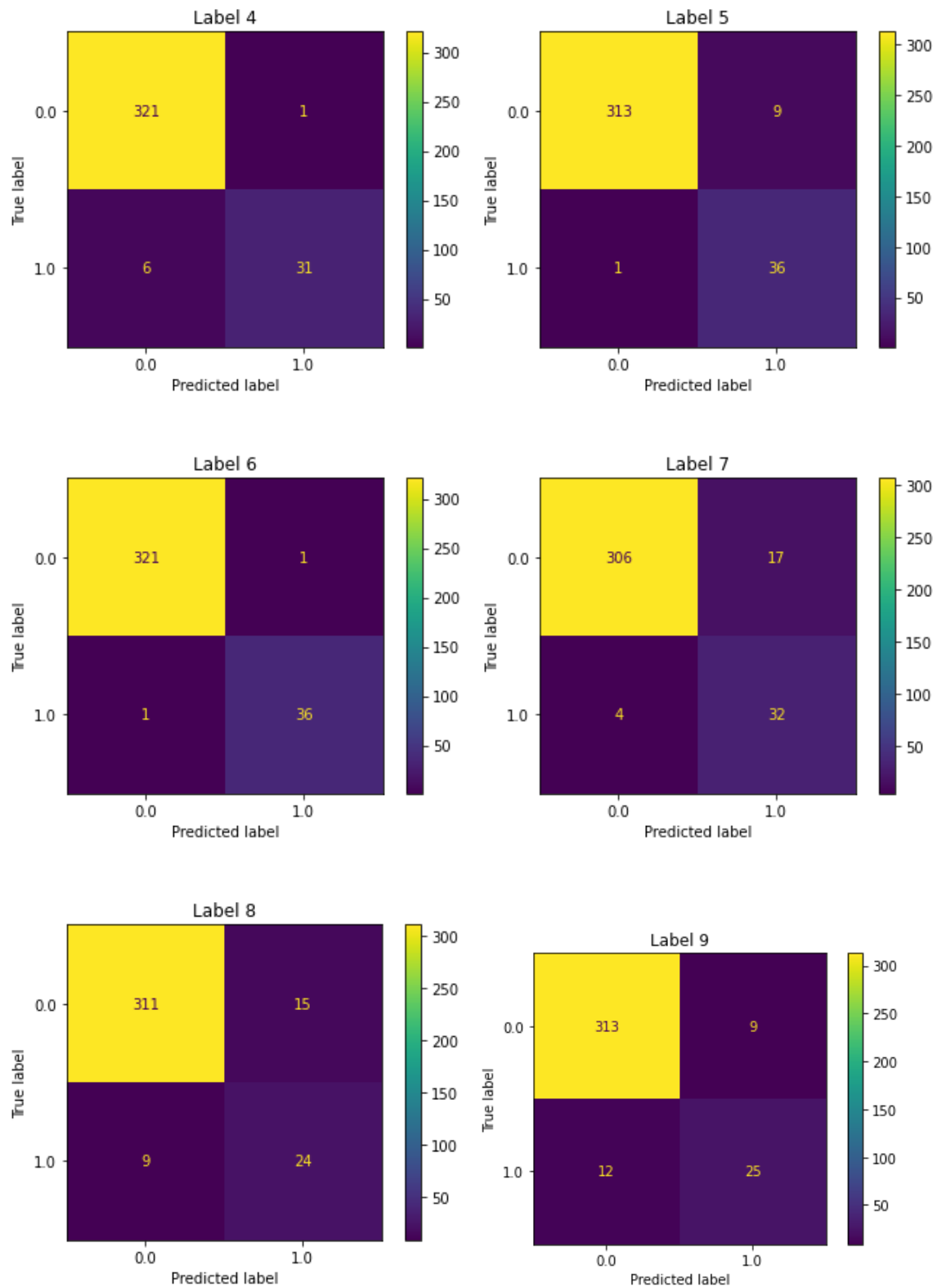
## Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)



**Naives Bayes (NB):** The accuracy is: 81.34 %.

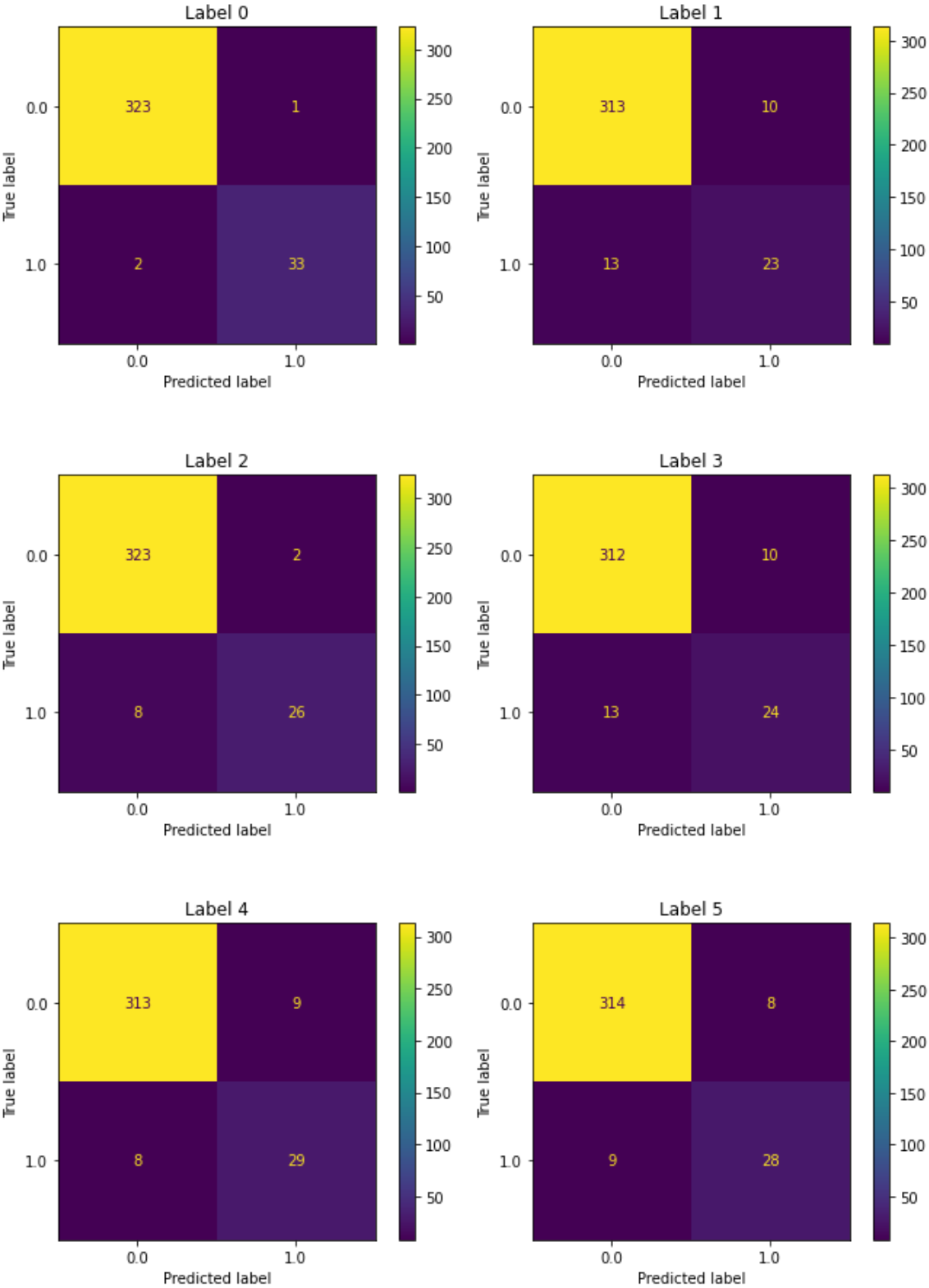


## Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)

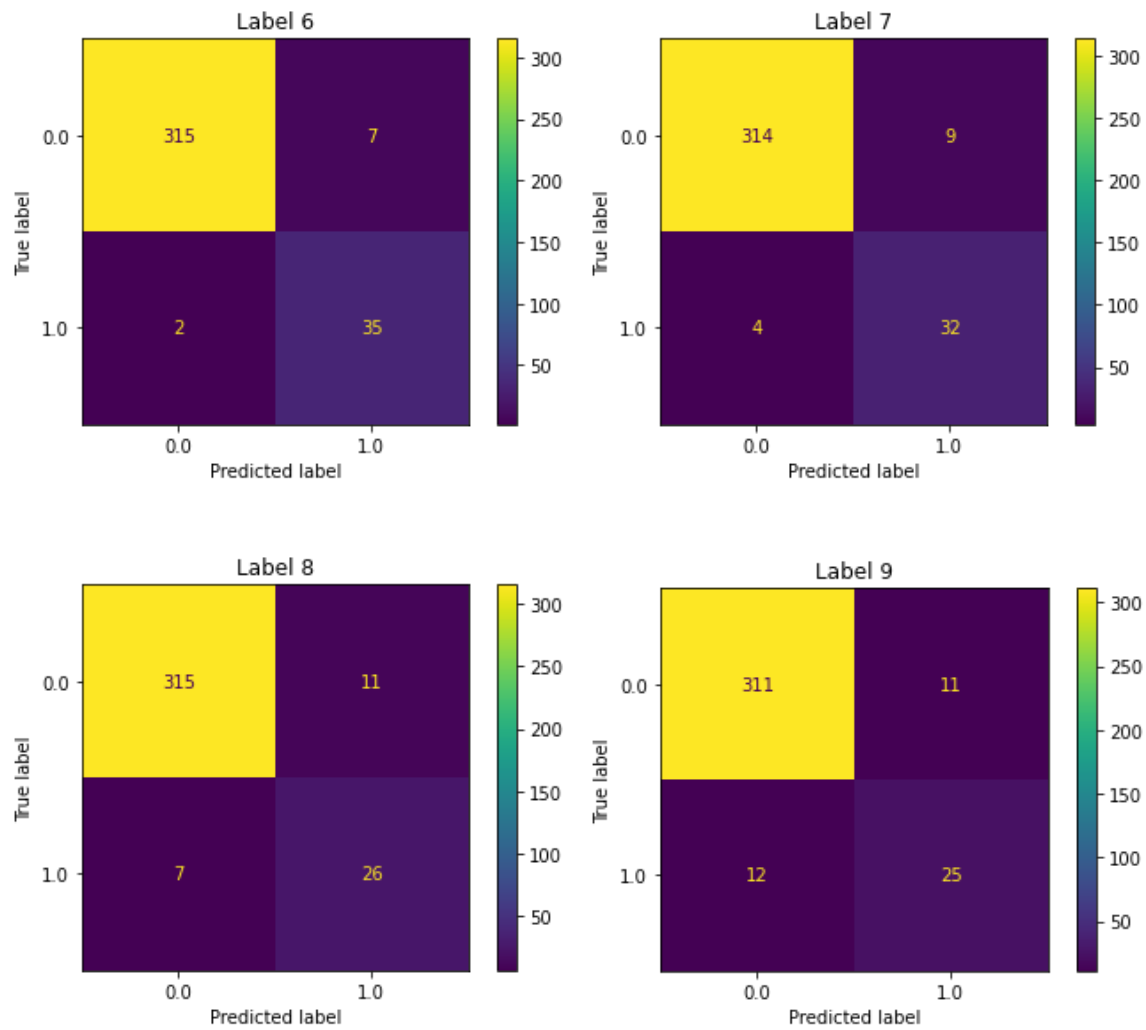


**CART METHOD:** The accuracy is: 78.27 %.

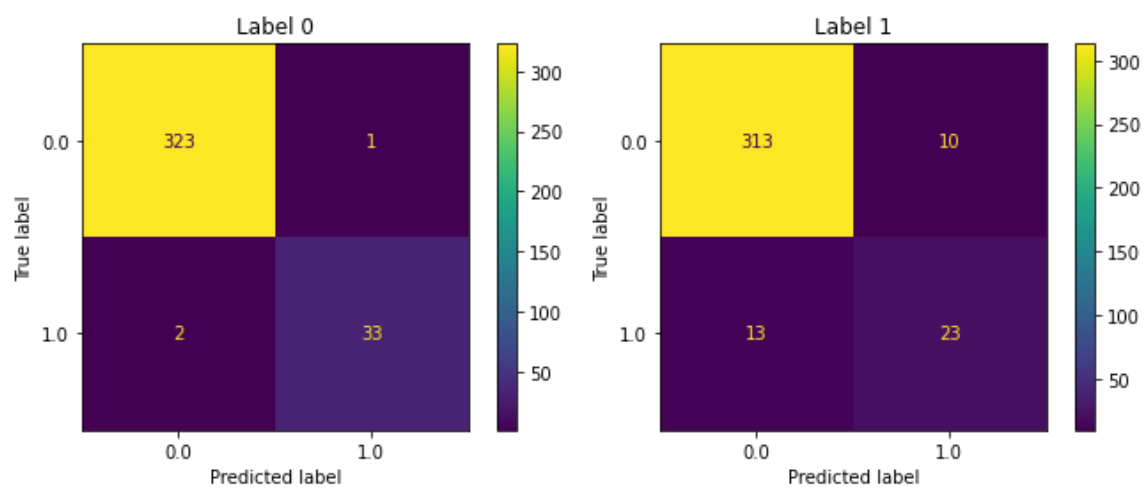
Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)



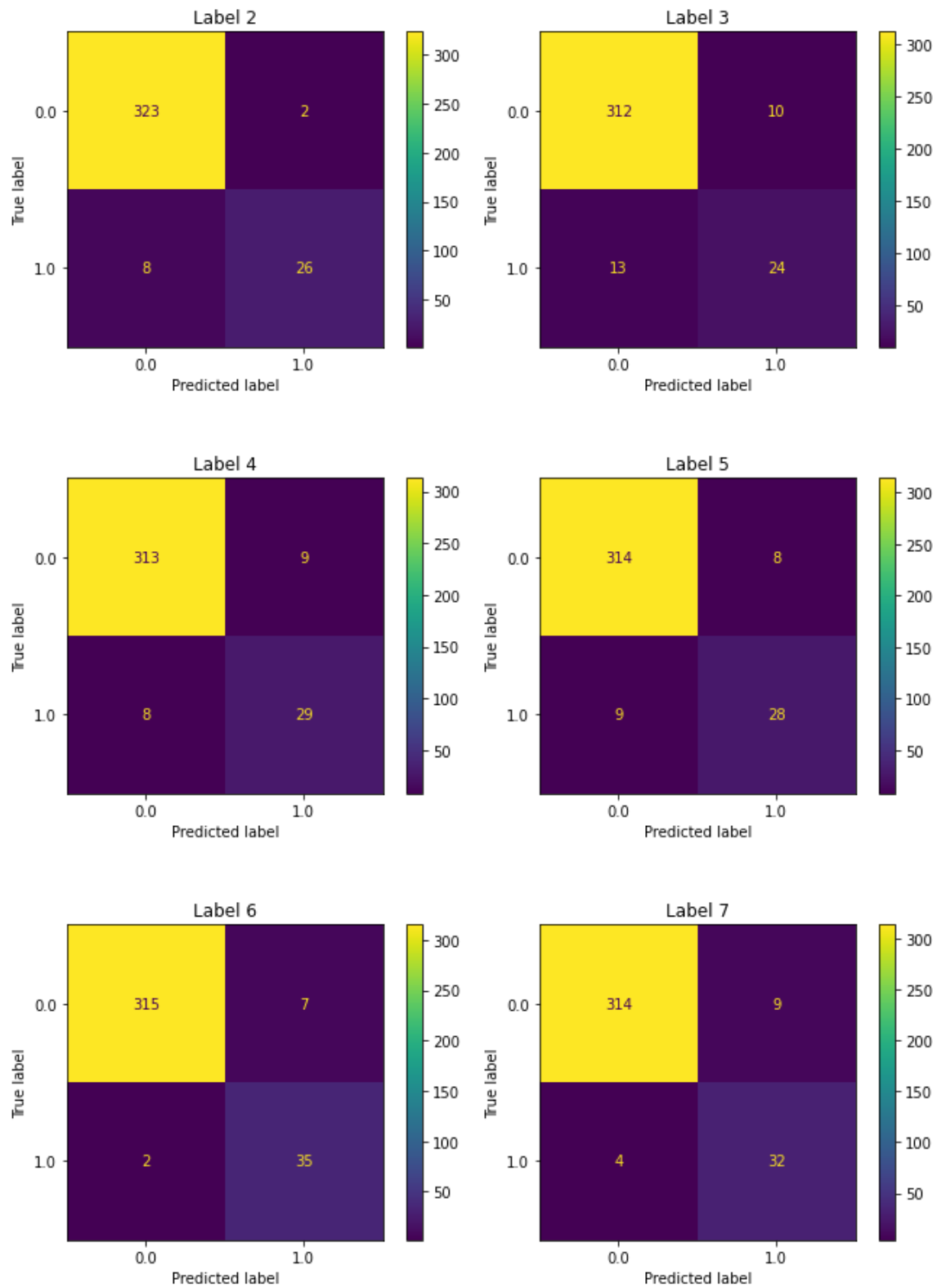
## Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)



**LDA:** The accuracy is: 78.27 %.

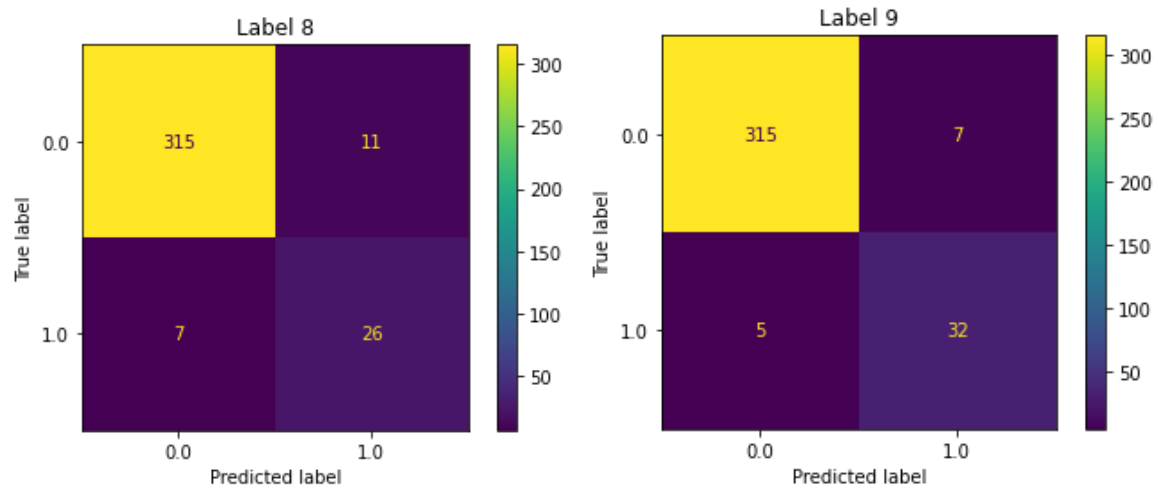


## Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)





## Procesamiento de imágenes médicas para el diagnóstico (Gpo 301)



El método que da el mejor resultado es el de KNN debido a que mostró una precisión del 96.66% al predecir las "labels", siguiendo de LDA con el 89.97%. los demás métodos presentan valores por debajo del 90% haciéndolos poco confiables para predecir la identificación de las imágenes que se están utilizando y muy posiblemente pueda ser confundido con otro tipo de imagen o en este caso confundirlo con otro de los dígitos que se evalúan.