Autocodificadores y la detección de anomalías

Los autocodificadores son capaces de aprender una representación eficiente de los datos de entrada. Este proceso se lleva a cabo mediante la creación de una red neuronal que transforma el espacio de entrada en un espacio latente y, posteriormente, lo reconstruye para obtener los datos originales a partir de esta representación reducida. Si los datos solo están compuestos por instancias de una sola clase, es posible entonces utilizar el autocodificador en contextos de clasificación de una sola clase y en la detección de anomalías. Veamos cómo se entrenaría para este último caso:

1. **Entrenamiento**. Durante esta fase se entrena el autocodificador con un conjunto de datos que solo contiene instancias de comportamiento normal. Una vez entrenado, si el autocodificador es capaz de reconstruir con precisión una entrada dada (es decir, la discrepancia entre la entrada original y la reconstruida es mínima), se infiere que la instancia pertenece al comportamiento normal dentro del espacio de datos. Sin embargo, si la discrepancia es significativa, existe una alta probabilidad de que la instancia sea considerada una anomalía.

2. **Establecimiento de un umbral de anomalía**. Se espera que los errores de reconstrucción asociados a los datos normales sean muy pequeños en comparación con el error de reconstrucción de una anomalía. Así, es posible establecer un umbral para determinar cuán grande debe ser la diferencia entre la entrada original y su reconstrucción para que una instancia sea considerada una anomalía.

3. **Detección**. Durante la fase de detección, cuando se presenta una nueva instancia al sistema, se calcula su diferencia de reconstrucción. Si esta diferencia supera el umbral previamente establecido, la instancia se marca como una anomalía (ver Fig. 2).

Fig. 2. Detección de anomalías con un autocodificador.

¿Cómo se puede establecer el umbral para la detección de anomalías? Existen varios métodos para hacerlo. A continuación, se exponen dos de ellos:

Enfoque estadístico. Una estrategia implica calcular la media y la desviación estándar del error de reconstrucción en el conjunto de entrenamiento. Luego, el umbral se define como un múltiplo fijo de la desviación estándar, como 2 o 3 veces por encima de la media. Las instancias cuyos errores de reconstrucción superen este umbral se clasifican como anomalías.

Validación cruzada. Si se dispone de datos etiquetados con anomalías, la validación cruzada puede ser útil para ajustar el umbral. Este proceso implica entrenar el autocodificador en un subconjunto de los datos y evaluar su rendimiento en otro. Ajustar el umbral basado en el rendimiento sobre un conjunto de validación puede ayudar a encontrar un punto óptimo que equilibre la sensibilidad y la especificidad de la detección de anomalías.

.

Bibliografía

Kalyan, S., Sridhar, A., (2023). *Beginning Anomaly Detection Using Python-Based Deep Learning: Implement Anomaly Detection Applications with Keras and PyTorch*. Apress.

Universidad de los Andes | Vigilada Mineducación Reconocimiento como
Universidad: Decreto 1297 del 30 de mayo de 1964.
Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.