

UNIVERSIDAD DEL VALLE DE GUATEMALA

Minería de datos

Sección 30

Ing. Leonel Guillen



Excelencia que trasciende

DEL VALLE
GRUPO EDUCATIVO

HT 7 - SVM

Bryan Carlos Roberto España Machorro - 21550

Javier Alejandro Prado Ramirez - 21486

Gabriel Garcia - 21352

Resultados y explicación:

En este informe se analizó un dataset en el cual veremos resultados y análisis de modelos de SVM en el cual el objetivo es poder predecir con ayuda de la data de entrenamiento ajustar el modelo para lograr con la data de prueba predecir con ayuda del modelo comparados con ayuda de distintos kernels en este caso se trabajó con kernels lineal, sigmoidal y polinomial, en el cual se comparó las Máquinas de SVM con Árboles de decisión.

Metodología:

Cargar Data: Se cargó la data del csv train y test para esta ocasión, y ver información importante para trabajar

Procesamiento:

- Identificar Variables numéricas y categóricas
- manejar valores nulos con imputer
- Escalar la data tanto test y train.

Modelos de clasificación evaluados:

1. SVM con kernel lineal
2. SVM con kernel sigmoide
3. SVM con kernel polinomial
4. Árboles de decisión

Resultados y Conclusiones

1. **SVM con kernel lineal:**
 - a. Precisión de 100%
 - b. La matriz de confusión predice con excelencia cada uno sin errores.
2. **SVM con kernel sigmoide:**
 - a. Precisión de 100%
 - b. La matriz de confusión predice con excelencia cada uno sin errores.
3. **SVM con kernel polinomial:**
 - a. Precisión de 68.9% aproximadamente
 - b. La matriz de confusión no acierta con precisión en las 3 categorías de precios de las casas pero es poco.
4. **Árboles de decisión:**
 - a. Precisión de 100% el modelo predice bien los resultados
 - b. La matriz de confusión revela que no se equivocó al predecir en las 3 categorías tanto 'Cheap', 'Medium', y 'Expensive'.

Análisis Comparativo

- **Mayor rendimiento VS peor rendimiento:** Vemos que 2 de 3 kernels lograron tener el 100 % estos son el lineal y sigmoide, dado que el polinomial fue el que tuvo una precisión muy baja de 68.9% aproximadamente, también al árbol de decisión le fue bien con precisión excelente.
- **KERNEL Lineal VS Sigmoide VS Polinomial:** Vemos una diferencia entre estos 3 , dado que el polinomio tuvo una gran diferencia de errores al predecir los precios de las clases de casas, en a comparación del lineal y sigmoide que sí lograron el 100% de precisión con el modelo

- **Conclusion:** El modelo con mejor rendimiento fueron los que utilizaron el kernel lineal y sigmoide dado que con estos lograron el 100% de precisión a comparación de polinomial

Recomendaciones

- Se recomienda usar kernel de sigmoides para lograr hacer predicciones multivariantes dado que este tiene la capacidad de procesamiento de 2 capas, por lo que rinde mejor que usar un kernel lineal.
- El Árbol de decisión es bueno, pero un poco tardado para las predicciones analizando cada uno de ellas y ordenandolas.