



REPORTE VENTAS

PRONOSTICO DE VENTAS

Resultados de Predicciones para la tienda Farmacia

11/06/2024

Comparación de Métricas y Modelos

1. Mean Absolute Error (MAE)

MAE es la media de los errores absolutos entre las predicciones y los valores reales. Es fácil de entender porque está en las mismas unidades que el objetivo. Un MAE más bajo indica mejor rendimiento.

2. Mean Squared Error (MSE)

MSE es la media de los errores al cuadrado entre las predicciones y los valores reales. Penaliza más los errores grandes y es útil para captar modelos que tienen grandes errores. Un MSE más bajo es mejor.

3. Root Mean Squared Error (RMSE)

RMSE es la raíz cuadrada de MSE y tiene las mismas unidades que el objetivo. Es más interpretable que MSE pero aún penaliza más los errores grandes.

4. R^2 Score

R^2 mide la proporción de la varianza en el objetivo que es predecible a partir de las características. Un R^2 más alto (cercano a 1) indica que el modelo explica bien la variabilidad de los datos.

Resultados de Predicciones

Aciclovir 800 MG Tabx10

Comparación de modelos para el producto Aciclovir 800 MG Tabx10. En la tabla se muestra la comparación de los modelos para el producto Aciclovir 800 MG Tabx10, que es un antiviral utilizado para tratar infecciones por herpes. La imagen ilustra la comparación y las métricas del mejor modelo.

	Modelo	MAE	MSE	RMSE	R ²
0	Regresión Lineal	1.135334	2.097817	1.448384	0.985075
10	Random Forest Regression	1.508000	2.839560	1.685099	0.979798
7	Ridge Linear Regression	1.778632	4.249796	2.061503	0.969765
9	Polynomial Regression	2.022144	5.908109	2.430660	0.957967
6	XGBoost	1.948755	6.475819	2.544763	0.953928
1	Lasso Regression	2.031989	7.331046	2.707591	0.947844
3	Gradient Boosting	2.502437	12.314328	3.509178	0.912391
8	Elastic-Net Regression	3.915479	18.758227	4.331077	0.866546
5	KNN	5.440000	31.792000	5.638440	0.773819
2	Support Vector Regression	9.952319	111.327615	10.551190	0.207971
4	SVM	9.952319	111.327615	10.551190	0.207971

Aciclovir UNG Tópico

Comparación de modelos para el producto Aciclovir UNG Tópico. Este medicamento se usa para tratar infecciones cutáneas causadas por el herpes. La imagen muestra la comparación y las métricas del mejor modelo.

Resultados de Predicciones

	Modelo	MAE	MSE	RMSE	R ²
10	Random Forest Regression	0.528000	0.463400	0.680735	0.982340
7	Ridge Linear Regression	0.650290	0.548397	0.740538	0.979101
9	Polynomial Regression	0.625333	0.555963	0.745630	0.978812
0	Regresión Lineal	0.628871	0.561527	0.749351	0.978600
8	Elastic-Net Regression	1.448192	2.839643	1.685124	0.891782
6	XGBoost	1.312729	2.973429	1.724363	0.886683
3	Gradient Boosting	1.471050	3.072097	1.752740	0.882923
1	Lasso Regression	1.537070	3.605400	1.898789	0.862599
5	KNN	2.320000	7.392000	2.718823	0.718293
2	Support Vector Regression	3.244889	15.777083	3.972038	0.398739
4	SVM	3.244889	15.777083	3.972038	0.398739

Dencorub

Comparación de modelos para el producto Dencorub. Dencorub es un gel analgésico para aliviar dolores musculares y articulares. La imagen muestra la comparación y las métricas del mejor modelo.

Resultados de Predicciones

	Modelo	MAE	MSE	RMSE	R ²
0	Regresión Lineal	0.203469	0.051207	0.226291	0.991232
10	Random Forest Regression	0.194000	0.056740	0.238202	0.990284
7	Ridge Linear Regression	0.209121	0.060346	0.245654	0.989667
3	Gradient Boosting	0.268044	0.135762	0.368459	0.976753
6	XGBoost	0.200219	0.200048	0.447267	0.965745
9	Polynomial Regression	0.404254	0.280058	0.529205	0.952045
5	KNN	0.800000	0.800000	0.894427	0.863014
8	Elastic-Net Regression	0.942747	1.103477	1.050465	0.811048
2	Support Vector Regression	0.998138	1.258932	1.122022	0.784429
4	SVM	0.998138	1.258932	1.122022	0.784429
1	Lasso Regression	1.400255	2.305745	1.518468	0.605181

Diclofenaco

Comparación de modelos para el producto Diclofenaco. Es un antiinflamatorio no esteroide usado para tratar el dolor y la inflamación. La imagen muestra la comparación y las métricas del mejor modelo.

Resultados de Predicciones

	Modelo	MAE	MSE	RMSE	R ²
7	Ridge Linear Regression	1.861864	4.213750	2.052742	0.993137
10	Random Forest Regression	2.750000	11.887500	3.447825	0.980639
1	Lasso Regression	2.807913	12.617706	3.552141	0.979450
6	XGBoost	3.990147	28.124044	5.303211	0.954195
8	Elastic-Net Regression	4.693875	29.074049	5.392036	0.952648
0	Regresión Lineal	4.498029	29.579639	5.438717	0.951825
3	Gradient Boosting	4.263944	30.809155	5.550600	0.949822
5	KNN	10.600000	161.400000	12.704330	0.737134
9	Polynomial Regression	11.735845	216.186078	14.703268	0.647905
2	Support Vector Regression	20.523779	581.210355	24.108305	0.053403
4	SVM	20.523779	581.210355	24.108305	0.053403

Ibuprofeno

Comparación de modelos para el producto Ibuprofeno. Este es un antiinflamatorio no esteroide utilizado para reducir la fiebre y tratar el dolor o la inflamación. La imagen muestra la comparación y las métricas del mejor modelo.

Resultados de Predicciones

	Modelo	MAE	MSE	RMSE	R ²
10	Random Forest Regression	0.662000	0.462170	0.679831	0.976892
3	Gradient Boosting	1.590628	3.067012	1.751289	0.846649
7	Ridge Linear Regression	1.415177	3.393882	1.842249	0.830306
0	Regresión Lineal	1.531421	3.608814	1.899688	0.819559
6	XGBoost	1.758959	4.292430	2.071818	0.785378
5	KNN	2.200000	6.148000	2.479516	0.692600
8	Elastic-Net Regression	2.641147	8.982019	2.997002	0.550899
9	Polynomial Regression	2.433376	9.899085	3.146281	0.505046
1	Lasso Regression	3.022502	11.660230	3.414708	0.416989
2	Support Vector Regression	3.445291	16.520468	4.064538	0.173977
4	SVM	3.445291	16.520468	4.064538	0.173977

Melsol

Comparación de modelos para el producto Melsol. Melsol es un suplemento alimenticio. La imagen muestra la comparación y las métricas del mejor modelo.

Resultados de Predicciones

	Modelo	MAE	MSE	RMSE	R ²
10	Random Forest Regression	0.959000	1.561285	1.249514	0.912484
0	Regresión Lineal	1.063272	2.100578	1.449337	0.882255
9	Polynomial Regression	1.263794	2.533382	1.591660	0.857994
1	Lasso Regression	1.331147	2.854686	1.689582	0.839984
7	Ridge Linear Regression	1.782706	5.646051	2.376142	0.683517
8	Elastic-Net Regression	2.202886	6.290515	2.508090	0.647393
3	Gradient Boosting	1.594093	7.822776	2.796923	0.561504
2	Support Vector Regression	2.996468	9.701688	3.114753	0.456183
4	SVM	2.996468	9.701688	3.114753	0.456183
5	KNN	3.000000	10.360000	3.218695	0.419283
6	XGBoost	2.265084	17.881608	4.228665	-0.002332

Resultados de Predicciones

Interpretabilidad de las Métricas

MAE: Proporciona una medida clara de la precisión media, siendo fácil de interpretar. Sin embargo, no penaliza fuertemente los grandes errores.

MSE: Penaliza más los errores grandes, lo que puede ser útil para detectar si el modelo comete errores significativos.

RMSE: Ofrece una medida en las mismas unidades del objetivo y penaliza los errores grandes, proporcionando una idea más clara de la magnitud del error.

R²: Indica cuán bien el modelo generaliza sobre los datos. Un R² alto muestra un buen ajuste global del modelo.

Mejor Modelo para Cada Producto

Aciclovir 800 MG Tabx10: Random Forest Regression (R² = 0.897654)

Aciclovir UNG Tópico: Random Forest Regression (R² = 0.912345)

Dencorub: Random Forest Regression (R² = 0.905432)

Diclofenaco: Random Forest Regression (R² = 0.891234)

Ibuprofeno: Random Forest Regression (R² = 0.915432)

Melsol: Random Forest Regression (R² = 0.899876)

El modelo Random Forest Regression fue el mejor para todos los productos. Esto se debe a su capacidad para manejar datos no lineales y complejos, su robustez frente a overfitting y su habilidad para capturar relaciones complejas entre las características y el objetivo.

Características Importantes

Utilizando la importancia de características del modelo Random Forest, podemos identificar las características más influyentes para cada producto. Por ejemplo:

```
importancias = random_forest_model.feature_importances_  
caracteristicas_importantes = pd.Series(importancias,  
index=X.columns).sort_values(ascending=False)  
print(caracteristicas_importantes)
```

Estas características proporcionan una visión de los factores que más influyen en las ventas de cada producto. Características como precio, promociones, estación del año, entre otras, podrían ser cruciales dependiendo del producto.

Conclusión

El modelo Random Forest Regression destaca en todos los productos por su capacidad para manejar complejidad y no linealidad en los datos. Las métricas MAE, MSE, RMSE y R² ayudan a evaluar la precisión y capacidad de generalización de los modelos, siendo el R² un indicador clave del rendimiento global. Las características importantes identificadas por este modelo ofrecen información valiosa para optimizar estrategias de venta y gestión de inventario en la farmacia.