




 Datasets:  aai510-group1 / telco-customer-churn   like 8

Follow  aai510-group1 3

Tasks:  Tabular Classification Modalities:  Tabular  Text

Formats:  csv Languages:  English Size: 1K - 10K

Tags:  tabular-classification  churn-prediction  telecom  customer-retention

 demographics  customer-service

Libraries:  Datasets  pandas  Croissant + 1

 Dataset card  Data Studio  Files  Community 1

Downloads last month 315

Use this dataset ▾



Size of downloaded dataset files:


1.89 MB

Size of the auto-converted Parquet files:

919 kB

Number of rows:

7,043

 Models trained or fine-tuned on aai510-group1/telco-customer-churn

 AIExplorer1234/customer-churn-prediction

Updated Feb 18

 Dataset Viewer

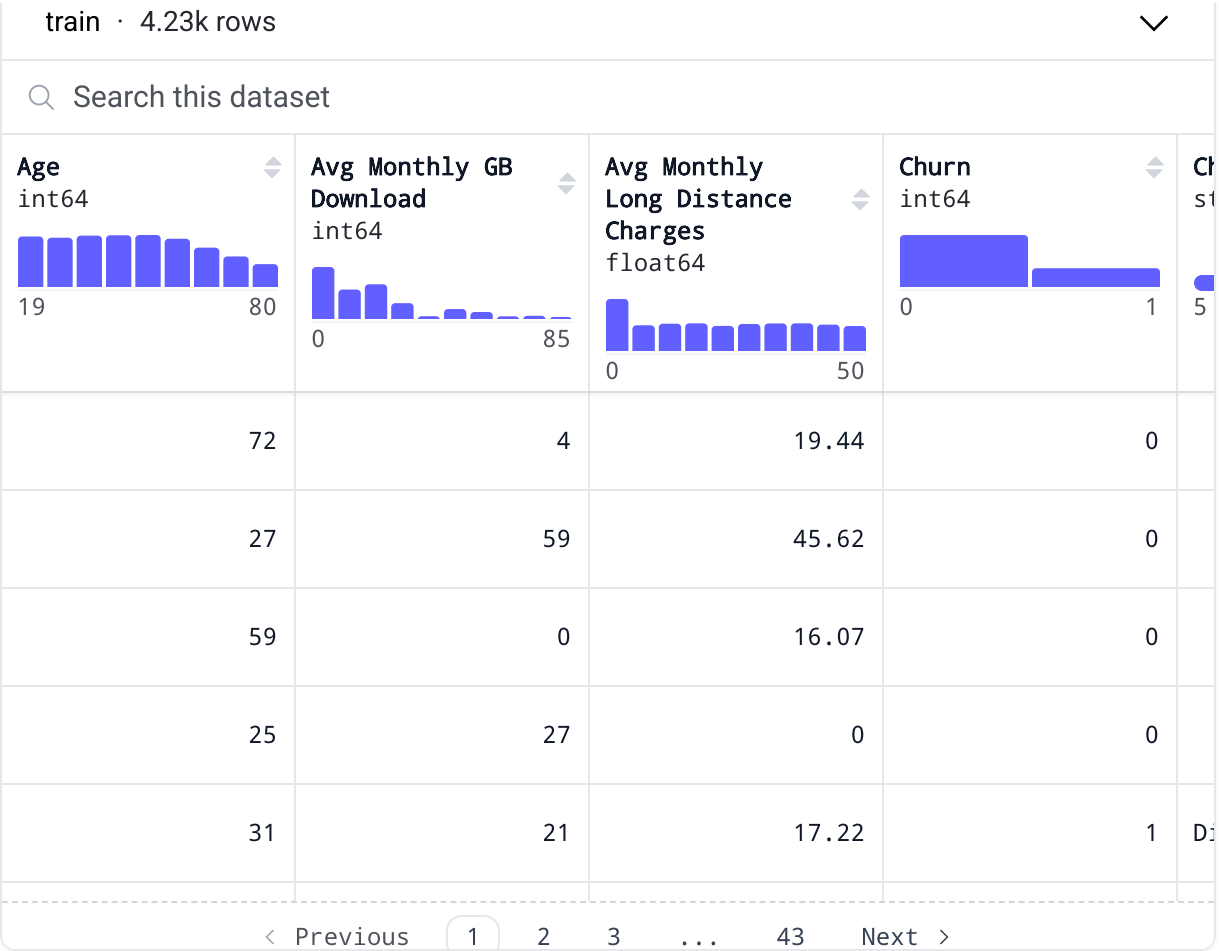
 Auto-converted to Parquet

 API

 Embed

 Data Studio

Split (3)



🔗 Dataset Card for Telco Customer Churn

This dataset contains information about customers of a fictional telecommunications company, including demographic information, services subscribed to, location details, and churn behavior. This merged dataset combines the information from the original Telco Customer Churn dataset with additional details.

🔗 Dataset Details

🔗 Dataset Description

This merged Telco Customer Churn dataset provides a comprehensive view of customer attributes, service usage, location data, and churn behavior. This expanded dataset is a valuable resource for understanding churn patterns, customer segmentation, and developing targeted marketing strategies.

[🔗 Uses](#)

[🔗 Direct Use](#)

This dataset can be used for various purposes, including:

- **Customer churn prediction:** Develop machine learning models to predict which customers are at risk of churning, leveraging the expanded features.
- **Customer segmentation:** Identify different customer segments based on demographics, service usage, location, and churn behavior.
- **Targeted marketing campaigns:** Develop targeted marketing campaigns to retain at-risk customers or attract new customers, tailoring campaigns based on the insights derived from the merged dataset.
- **Location-based analysis:** Analyze customer churn trends based on specific locations, cities, or zip codes, and identify potential regional differences.

[🔗 Out-of-Scope Use](#)

The dataset is not suitable for:

- **Real-time churn prediction:** The dataset lacks real-time data, making it inappropriate for immediate churn prediction.
- **Personal identification:** While the dataset contains customer information, it is anonymized and should not be used to identify individuals.

[🔗 Dataset Structure](#)

The dataset is structured as a CSV file with 49 columns, each representing a customer attribute. The columns include:

- **Age:** The customer's age in years.

- **Avg Monthly GB Download:** The customer's average monthly gigabyte download volume.
- **Avg Monthly Long Distance Charges:** The customer's average monthly long distance charges.
- **Churn Category:** A high-level category for the customer's reason for churning.
- **Churn Label:** Indicates whether the customer churned.
- **Churn Reason:** The customer's specific reason for leaving the company.
- **Churn Score:** A score from 0-100 indicating the likelihood of the customer churning.
- **Churn Value:** A numerical value representing whether the customer churned (1 for churned, 0 for not churned).
- **City:** The city of the customer's residence.
- **CLTV:** Customer Lifetime Value.
- **Contract:** The customer's contract type.
- **Country:** The country of the customer's residence.
- **Customer ID:** A unique identifier for each customer.
- **Customer Status:** The customer's status at the end of the quarter (Churned, Stayed, or Joined).
- **Dependents:** Whether the customer has dependents.
- **Device Protection Plan:** Whether the customer has a device protection plan.
- **Gender:** The customer's gender.
- **Internet Service:** Indicates whether the customer subscribes to internet service.
- **Internet Type:** The type of internet service provider.
- **Lat Long:** The combined latitude and longitude of the customer's residence.

- Latitude: The latitude of the customer's residence.
- Longitude: The longitude of the customer's residence.
- Married: Indicates if the customer is married.
- Monthly Charge: The customer's total monthly charge for all their services.
- Multiple Lines: Whether the customer has multiple phone lines.
- Number of Dependents: The number of dependents the customer has.
- Number of Referrals: The number of referrals made by the customer.
- Offer: The last marketing offer the customer accepted.
- Online Backup: Whether the customer has online backup service.
- Online Security: Whether the customer has online security service.
- Paperless Billing: Whether the customer has paperless billing.
- Partner: Whether the customer has a partner.
- Payment Method: The customer's payment method.
- Phone Service: Whether the customer has phone service.
- Population: The estimated population of the customer's zip code.
- Premium Tech Support: Whether the customer has premium tech support.
- Quarter: The fiscal quarter for the data.
- Referred a Friend: Indicates if the customer has referred a friend.
- Satisfaction Score: The customer's satisfaction rating.
- Senior Citizen: Whether the customer is a senior citizen.
- State: The state of the customer's residence.
- Streaming Movies: Whether the customer has streaming movies service.
- Streaming Music: Whether the customer has streaming music service.
- Streaming TV: Whether the customer has streaming TV service.

- **Tenure in Months:** The number of months the customer has been with the company.
- **Total Charges:** The customer's total charges.
- **Total Extra Data Charges:** The total charges for extra data downloads.
- **Total Long Distance Charges:** The total charges for long distance calls.
- **Total Refunds:** The total refunds received by the customer.
- **Total Revenue:** The total revenue generated by the customer.
- **Under 30:** Indicates if the customer is under 30 years old.
- **Unlimited Data:** Whether the customer has unlimited data.
- **Zip Code:** The zip code of the customer's residence.

[🔗 Dataset Creation](#)

[🔗 Curation Rationale](#)

This merged dataset was created to provide a more comprehensive and detailed analysis of customer churn behavior. Combining multiple sources of data allows for a richer understanding of factors influencing churn.

[🔗 Source Data](#)

[🔗 Data Collection and Processing](#)

The dataset is derived from the original Telco Customer Churn dataset and additional data sources. The specific data collection and processing methods are not disclosed.

[🔗 Bias, Risks, and Limitations](#)

[🔗 Bias](#)

The dataset may exhibit biases due to the simulated nature of the original Telco Customer Churn data. It is essential to consider that the dataset may not accurately reflect the demographics, service usage, or churn patterns of actual telecommunications companies.

Risks

Using the dataset for real-world decisions without proper validation and understanding of its limitations can lead to inaccurate predictions and potentially biased outcomes.


Limitations

- **Simulated Data:** The dataset is based on simulated data and may not fully represent real-world customer behavior.
- **Limited Context:** The dataset may lack specific contextual information such as customer feedback or reasons for churn.
- **Potential Bias:** The simulated data may not fully capture the nuances of customer behavior and churn patterns, especially when combined with additional data sources.

Recommendations

Users should be aware of the dataset's limitations and potential biases. Consider the following:

- **Validation:** Validate the dataset's results against real-world data before making critical decisions.
- **Contextualization:** Include additional contextual information if available to improve model accuracy and insights.
- **Transparency:** Be transparent about the dataset's limitations and potential biases when communicating results.

 System theme

Company

TOS

Privacy

About

Jobs

Website

Models

Datasets

Spaces

Pricing

Docs

