July 2021

Inferring Restaurant Affordability From Menu Prices

Xinru Yang

Inferring Restaurant Affordability From Menu Prices

ABSTRACT

Online information sources such as websites, search engines, digital maps, etc. provide restaurant information which often includes an affordability indicator that indicates the general level of prices charged by the restaurant for the items on its menu. Currently, such affordability indicators are derived from subjective and expensive consumer surveys, which can be prone to error.

This disclosure describes techniques to automatically infer an affordability indicator for a restaurant using data obtained from existing data sources and provided as input to a suitably trained machine learning model. Prices are obtained from sources of high-quality structured information and/or lower-quality unstructured information, such as user-generated content. A trained model utilizes the obtained price information and generates the affordability indicator. The model is designed to take into account local differences in pricing and cost of living via appropriate normalization that enables the operation to scale globally in a seamless manner.

KEYWORDS

- Price indicator
- Affordability indicator
- Restaurant affordability
- Price inference
- Restaurant rating
- Food ordering
- Restaurant menu

BACKGROUND

Affordability is an important criterion when choosing a restaurant for dining. Sites and services that provide information about restaurants typically include an indicator of affordability in terms of the overall level of prices charged by the restaurant for the items on its menu. The affordability indicator is typically presented in the form of a rating using currency symbols. For instance, restaurants may be rated with $, $$, $$$, or $$$$ to mark them as inexpensive, moderately priced, expensive, and very expensive, respectively. As such, the affordability indicator is distinct from other pricing information such as ‘typical price’ (which is a specific currency value) or ‘price range’ (which is a low-high range of prices).

Currently, the affordability indicator is derived mainly from consumer surveys that ask questions to people who have been to a restaurant. For instance, consumers can be asked to assign a price-level indicator for restaurants and the weighted average of these votes can be used to set the dividing lines between the affordability-indicator buckets. Apart from being expensive and time-consuming, data collected from consumer self-reports is subjective and can have inaccuracies due to limitations of human memory. Moreover, the low conversion rate for survey requests can result in uneven coverage across restaurants, leaving some restaurants without enough data to derive an affordability indicator.

DESCRIPTION

This disclosure describes techniques to automatically infer an affordability indicator for a restaurant using data obtained from existing data sources and provided as input to a suitably trained machine learning model. Specifically, information for the prices of the various items offered in the menu of a restaurant is obtained using high-quality structured information on prices, e.g., via an application programming interface (API) that includes the option to obtain prices, via online food ordering services (with permission from the respective service provider), etc.

In case menu data is not available in such a structured form for certain restaurants, corresponding unstructured data can be obtained via alternate mechanisms, such as from the restaurant website; user-generated content such as menu photos, posts or comments regarding prices, etc. accessed with permission; etc. Structured pricing information is extracted from the unstructured data by employing appropriate techniques, such as optical character recognition (OCR) to derive text information contained in photos followed by extracting menu items and prices from the text, some or all of which may be performed by a trained machine learning model.

The data obtained on prices for restaurants is combined with additional information regarding the restaurant, such as type of cuisine, location, etc. The restaurant information and pricing data is provided as input to a trained model that outputs the appropriate affordability indicator for the restaurant based on the input. The automatically derived indicator is then displayed along with other restaurant information shown to users.

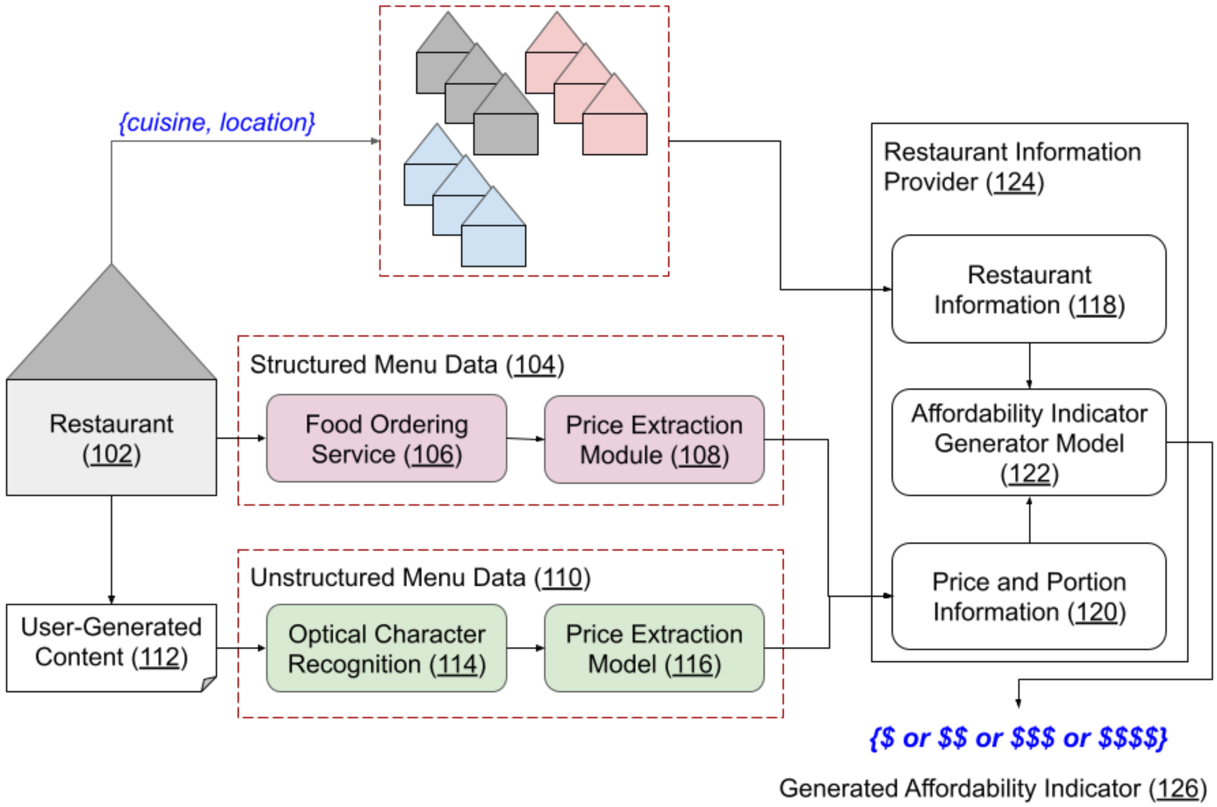

Fig. 1: Generating affordability indicator for a restaurant by extracting menu prices

Fig. 1 shows an example of operational implementation of generating affordability indicators for restaurants, per the techniques described in this disclosure. Data about the menu of a restaurant (102) is obtained in structured form (104) from a food ordering service (106) or other structured data source, with appropriate permission, and used to extract information on prices and portions, e.g., by a price extraction module (108, for the menu items.

Alternatively, or addition, user-generated content (112) about the restaurants is used with permission to obtain menu data in unstructured form (110). Information on prices and portions is obtained from the unstructured data by a machine learning model (116) trained to extract the information from text derived via OCR (114). The extracted price and portion information (120) and other relevant restaurant information (118), such as cuisine, location, etc., is aggregated by a restaurant information provider (124) and provided as input to a trained machine learning model

(122). The model generates an affordability indicator (126) for the restaurant which may be provided to users via an appropriate user interface, e.g., in query responses from a search engine or virtual assistant, in a digital maps interface, etc.

The machine learning model is designed to take into account local differences in pricing and cost of living. For instance, a dish can be priced higher in a large city compared to a small town in a rural area. Similarly, the model can account for differences in baseline prices based on the type of cuisine. For instance, some cuisines may be inherently more expensive than others. Additionally, price can depend on the size of the portion in terms of weight, dimensions, number of pieces, etc. To that end, calculation of the indicator can include computing the average price across all menu items for a restaurant and normalizing it for the region, cuisine type, and portion (if applicable). The normalized data can then be aggregated to train, use, and evaluate the model.

The model that generates the affordability indicator can be any suitable machine learning model, such as a neural network; a regression model based on relevant metrics such as mean-squared error (MSE), root mean-squared error (RMSE), etc.; a classification model; etc. The model can be trained and evaluated using labeled data such as that obtained from consumer surveys that are used to derive the current affordability indicators. Further, model performance using noisy, lower-quality unstructured data can be compared to that using high-quality structured data to ensure that using the lower-quality data yields acceptable results.

The affordability indicator for a restaurant can be updated by re-running the model, e.g., whenever there is a change in the menu data for the restaurant. Moreover, with permission, the model can be improved by incorporating consumer feedback on the accuracy of the derived affordability indicator. Alternatively, or in addition, human raters can visit restaurants and rate the affordability of their prices, especially in the cases of restaurants that are close to the thresholds for different affordability levels.

The techniques described in this disclosure enable generation of an inferred affordability indicator for a restaurant based only on aggregated menu information. Such an approach provides greater coverage across restaurants and is faster and cheaper than one that relies on consumer surveys. The techniques can be implemented by any platform or service that aggregates and provides restaurant information, such as search engines, virtual assistants, review websites, digital map applications, etc. Normalization based on appropriate parameters, such as region, type of cuisine, etc., can enable the operation to scale globally in a seamless manner.

In addition to restaurants, the techniques can be applied to generate price-level indicators for other types of businesses and services, such as hair salons, stores, etc., as long as relevant pricing information of their offerings can be obtained.

CONCLUSION

This disclosure describes techniques to automatically infer an affordability indicator for a restaurant using data obtained from existing data sources and provided as input to a suitably trained machine learning model. Prices are obtained from sources of high-quality structured information and/or lower-quality unstructured information, such as user-generated content. A trained model utilizes the obtained price information and generates the affordability indicator. The model is designed to take into account local differences in pricing and cost of living via appropriate normalization that enables the operation to scale globally in a seamless manner.