

Case Study

Hotel Analysis Platform by AltexSoft

AltexSoft Builds Machine Learning Driven Platform to Score Hotel Amenities
Based on Customer Reviews

Python, NLTK, Keras, TensorFlow, JavaScript

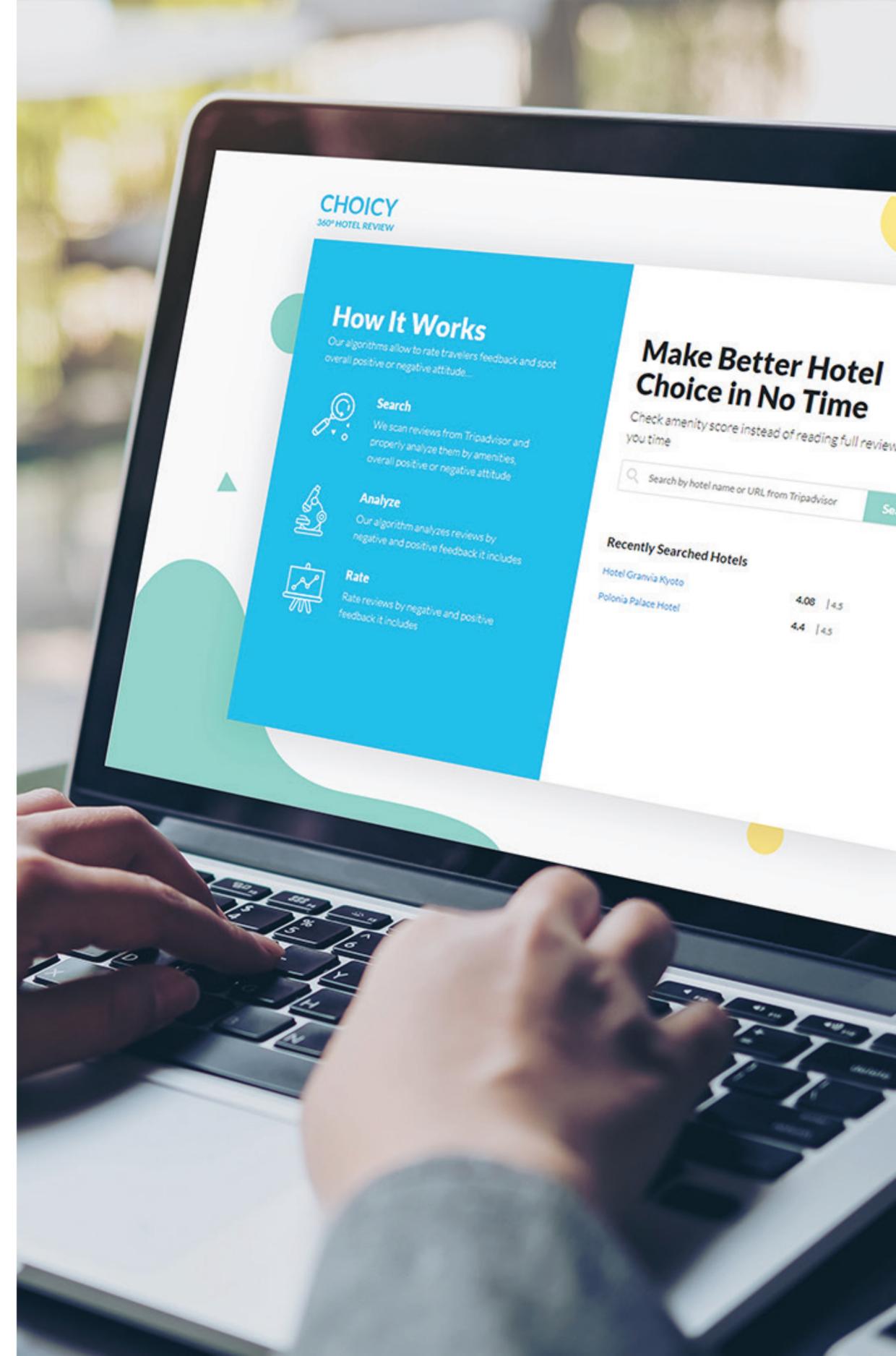


Background

Travel industry uses sentiment analysis to understand customers and improve their experience. Brand monitoring, competitive research, product analysis – these are just several ways OTAs and hotels apply it. The AltexSoft data science team considered a traveler-facing spin on this technology.

Choicy is a natural language processing (NLP) application that can help travelers get the most detailed information about a hotel, check amenities ratings, and compare hotels based on these parameters. The tool aggregates customer-created hotel reviews from public sources, analyzes them, and then generates amenity quality ratings for each hotel.

Choicy, started as an internal proof of concept, is now ready for integration into travel review platforms and online travel agencies that distribute hotels.



Challenges

To create an application, our team faced the following tasks:

1.

Define a detailed concept of the product

2.

Prepare a training dataset

3.

Find a suitable NLP model to score the amenities

4.

Design and develop the application UI

Value Delivered

1. Defining hotel amenities comparison criteria and the overall product concept

As the idea emerged and product concept was discussed, our team outlined the product workflow as follows: A user searches for a hotel by name or URL. The system detects amenities descriptions in reviews, including those discussing bathroom, lounge, breakfast, air conditioning, location, etc. Then a trained algorithm analyses the reviews for word-markers of positive, neutral, or negative guest experiences of the amenities. The results for each category and an overall score of a hotel are aggregated and returned to a user. As a result, a user can explore individual amenity ratings and compare them to the other hotels without reading full reviews. Since the implementation needed machine learning, our team started product development with training dataset preparation.

3. Training a model to score and analyze the reviews by amenities

Our engineers tried several different approaches to find a proper model to score the amenities, like classic natural language processing approaches and the latest deep learning models like BERT. As a result, they decided on two neural networks for sentiment analysis, a convolutional neural network for scoring hotels by whole reviews (1D-CNN with GloVe embeddings) and HAPN (hierarchical attention based position-aware network) to score individual amenities. This allows users to get both general impressions about hotels and drill down into specific service details if those details are critical for a pleasant stay. The final step of Choicy development was the user interface design and engineering.

2. Preparing a dataset with labeled sentiment for amenities

For model training, we collected a hotel review dataset that consisted of about 100,000 samples from public sources, including datasets available at Kaggle. As these datasets didn't have labeled sentiment for various hotel amenities, the team had to label them semi-manually. Once the dataset was ready, data scientists could proceed to model training.

4. Creating a user-friendly interface

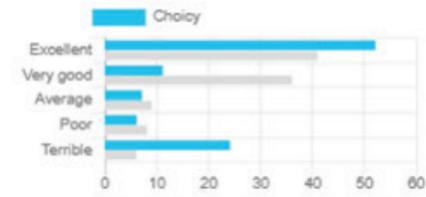
Our UX/UI designer suggested a minimalistic design of the application interface. A home page and the page with the reviews has a search box where a user can type a hotel's name or insert a URL. The page has various navigation components like tags that help simplify the search, compare the hotels by the amenities, and categorize them. The user interface was created using JavaScript.

Novotel Paris Centre Gare Montparnasse

17 rue du Cotentin, 75015 Paris France

Compare to Another Hotel

Overall Score What's the difference?



Amenity Score How does it work?

What was really good?

Bar/Lounge	●●●●●	21
Business Trip	●●●●●	16
Cleanliness	●●●●●	47
Comfort	●●●●●	33
Food	●●●●●	60
Fridge	●●●●●	10
Location	●●●●●	49
Quietness	●●●●●	9
Service	●●●●●	72

What was OK?

Bathroom	●●●●●	25
Coffee	●●●●●	12
Room	●●●●●	76
Sleeping	●●●●●	23
Value	●●●●●	22

What was very bad?

Air Conditioner	●●●●●	7
-----------------	-------	---

Reviews

room

bathroom X

room X

Top positive reviews Show more

Amenity Score ●●●●●

Our rooms were very comfortable, great walk in shower, memory foam mattress and lots of good quality pillows.

Drib63 Hertfordshire, United Kingdom

Show Full Review

Amenity Score ●●●●●

The rooms were big and spacious and loved the fact that the rooms have a two room bathroom so that families can easier get ready.

josiemcnally Vicenza, Italy

Show Full Review

Amenity Score ●●●●●

Apparently that wasn't possible so we never did get clean sheets or clean towels during our stay.

Gemma O

Show Full Review

show more...

Top negative reviews Show more

Amenity Score ●●●●●

Showers were weak, not particularly hot, and leaky (on the floor). WC or toilet was in a separate room from the sink and shower. Sinks were super shallow without a soap dish.

2C2_travel Boston, Massachusetts

Show Full Review

Approach and Technical Info

The project's scope was **7.5 man-months**. The product was completed over the course of 2 months by a team of five professionals: **2 machine learning engineers**, **a software engineer**, **a UI/UX designer**, and **a project manager**.

The technology stack of the project included **Python**, **NLTK**, **Keras**, and **TensorFlow** (for the algorithms), **HTML** and **JavaScript** (for the user interface).

Hotel Analysis Platform by AltexSoft

AltexSoft US Sales HQ

701 Palomar Airport Road, Suite 300,
Carlsbad, CA 92011
+1 (877) 777-90-97

AltexSoft Global HQ

32 Pushkinskaya Str.,
Kharkiv, Ukraine 61057
+38 (057) 714-1537
sales@altexsoft.com