white paper

Intellexer™

Sentiment Analyzer SDK

Identifying the opinions, attitudes and emotions of peoples
Understanding the voice of the customers
Enhancing enterprise systems

Incorporated modules:
Preformator, Linguistic Processor

www.intellexer.com

© 2014 EffectiveSoft, Ltd.

What is Sentiment Analysis?

Sentiment Analysis can answer to the question – *What people think about my product/brand/service?*

Sentiment analysis refers to the use of natural language processing and computational linguistics to identify and extract subjective information in text documents: opinion polarity (positive/negative/neutral) with its holder, objects or product features with associated sentiment phrases and affective states (emotions) of the opinion holders.

At present due to the intense increase of opinionated texts like product review websites, tweets and blog posts, the development of systems that provide solution for a wide range of sentiment analysis problems becomes more essential.

There are two main approaches for sentiment analysis: lexicon-based and supervised methods. The lexicon-based approach calculates the semantic orientation of words in a text by obtaining word polarities from a lexicon. Supervised learning approaches use statistical machine learning techniques to establish a model from a large corpus of documents. The set of sample opinions form the training data from which the model is built. Applying machine learning techniques can achieve over 80% in sentiment accuracy. Unfortunately, this level of accuracy is not quite on par with the performance for full opinion text understanding.

The main factors that makes sentiment analysis difficult are that the structure of opinions is often context-sensitive and domain dependent.

Why Intellexer Sentiment Analyzer SDK?

Extracts sentiments, opinions and emotions

Tracks sentiment for services, products and brands

Intellexer Sentiment Analyzer SDK is the powerful and efficient solution which automatically detect sentiment (positivity/negativity), opinion objects and emotions (liking, anger, disgust and etc.) from unstructured text information.

The team of R&D department of EffectiveSoft Ltd. has developed a unique hybrid approach to sentiment analysis which, unlike existing methods, is based on using not only linguistic and statistical information, but also the set of complex semantic rules. This technique provides the one of the best sentiment accuracy results on the market and the system adaptation to the knowledge domain.

 © 2014 EffectiveSoft, Ltd.

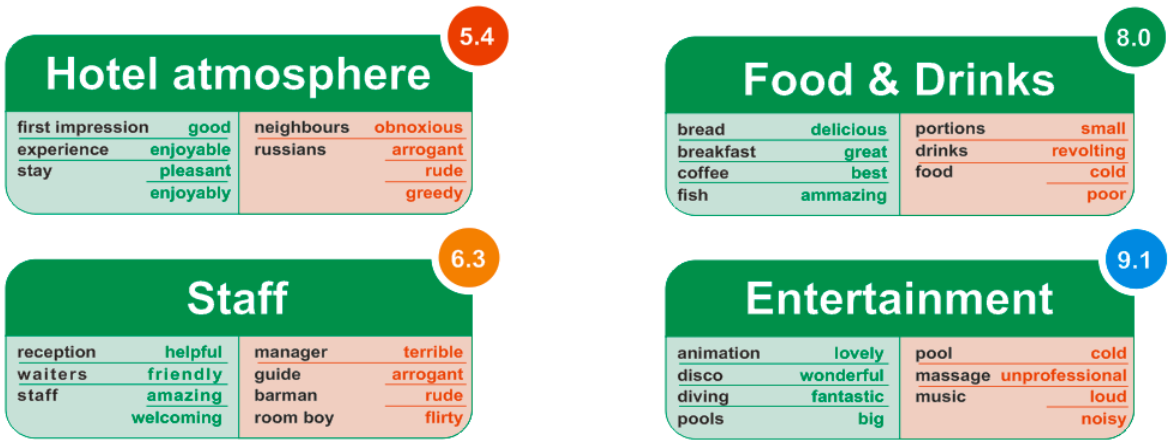
What can Intellexer Sentiment Analyzer SDK do?

Extracts sentiments, opinions and emotions

Our approach to sentiment analysis successfully process:

- Context-depended opinions (short battery life → short recharge time)
- Domain-depended opinions (small memory size (Gadgets) → small price size (Restaurants))
- Negatives and Intensifiers (not a bad service, too cheap device)
- Connectors (Although this hotel has brilliant location, it has a horrible service – negative opinion)
- Non-opinion phrases containing opinion words (romantic prose, stained glass)

Our experts create custom built ontologies for the most significant and widely used data domains (hotels, restaurants, gadgets and etc.). In order to obtain informative reports opinion objects can be grouped in different ways; it depends on a client's preferences.



 © 2014 EffectiveSoft, Ltd.

Document Sentiment Classification

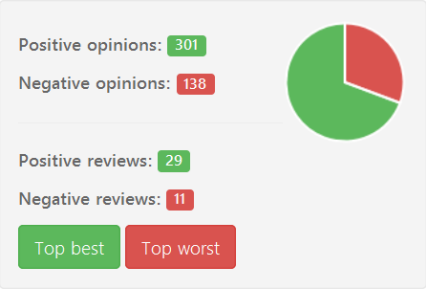
Determine whether a review is positive or negative

Document Sentiment Classification aims to classify a whole document text as expressing a positive, neutral or negative opinion. This task would be helpful in recommender systems and business intelligence software, where customer negative or positive feedback could be quickly find.

Intellexer Sentiment Analyzer uses a wide range of text features for document sentiment classification:

- Part of speech tags
- Opinion words and syntactic dependency between them
- Opinion objects with associated sentiment phrases
- Positions of sentiment sentences through the document text

For our experiments, we create the dataset which consists of more than 250 000 hotel and restaurant reviews from TripAdvisor.com divided into positive (4 and 5 stars user rating reviews) and negative (1 and 2 stars user rating reviews) categories. At this collection we achieved sentiment classification accuracy over 91% (typical competitors result: 80–85% accuracy).



Top Best Reviews

+13.08 Christopher K. 30.09.2013

My family and I (6 of us) had reservations for 8:30 last Saturday.
We were all a little nervous when we got closer to the place and it looked like a beat up building.
Once we opened the door though, we were happy to see it was a bustling place with a really great atmosphere.
We were seated immediately and the service was excellent throughout the night.
I wish I remembered the server 's name, she was really friendly, fast and knowledgeable.
We had calamari, bon chon wings and clams for starters.
Calamari was a little chewy, but the wings and clams were tasty.
Our cocktails were excellent.
I had the Bijou, which is gin-based.
Other people mixed it up and all were happy with their choices.
Overall our entrees were great.
Chicken schnitzel was the best I 've ever had (even better than the one I had in Vienna).
Others got mac n cheese, ahi tuna and sirloin.
Everyone really loved it.
The only gripe would be that the steak was ordered medium rare and came back almost well done.
The flavor was there but it was definitely over-cooked.
Considering how busy it was, we left incredibly happy with our experience.
I will be bringing people back here in the future.
Great first experience!

Top Worst Reviews

-11.92 Linda Y. 20.11.2013

Everyone should note that the high rating this place gets is for the drinks, not the food.
I 've had a few cocktails here, and I 've enjoyed them all thoroughly.
Their cocktail list is never ending, so if I were you I would just look at the short list instead of the full menu.
This place has a loud atmosphere, so it 's great for a drinking get together (though the bar space is very tight)
The entrees were overpriced, and though some of the ingredients were used in interesting ways.
Appetizers :
- gruyere fritters - these are the modern classier cousin of the jalapeno popper.
It was served over this creamy heavy white sauce, which was delicious, but definitely made me a bit too full at the end.
- grilled calamari - beautifully presented, but unfortunately undercooked.
Entrees :
- duck confit - though this was an appetizer, I ordered it as an entree because I found all their main courses overpriced.
The chef should take a few lessons because duck leg was overcooked.
I would have complained, but it was so dark that I could not tell if it was perfectly seared or perfectly burnt.
Either way, the meat came out way too tough for a confit.
Beside the occasional bits of salt, they were completely flavorless.
whoever did the prep should learn how to toss the overgrown thick stems.
It was more sticky and the honey in the ice cream or the cake made it almost sickeningly sweet.
- decaffeinated coffee - talk about stale coffee.
It was clearly sitting on the burner for way too long.

© 2014 EffectiveSoft, Ltd.

Review Rating Prediction

Understanding the customer feedback and real opinions rating

Review Rating Prediction aims to predict a user's numeric rating (usually from 1 to 5 stars) in a review from the text of the review. Rating Prediction task can be effectively used to:

- Help web users to determine the rating of their reviews
- Detect suspicious or fake online reviews
- Better understanding of customer opinions
- Improve the default (5 star) rating system or create proprietary ranking scale for more precise user's feedback representation

Review Rating Prediction has traditionally been model as

a multi-class classification or regression task. This techniques can achieve 70-78% in rating prediction accuracy.

Our prediction algorithm is based on vector space model. In this model, user reviews are represented as vectors in multidimensional space, whose components are opinion words, intensifiers and opinion objects with associated sentiment phrases. We apply the improved term frequency/inverse document frequency approach which uses semantic features of opinion words to enhance the accuracy of rating prediction. The similarity between user review and rating category is determined as cosine angle between appropriate vectors.

To evaluate the effectiveness of our prediction algorithm we create a dataset of hotel and restaurant reviews from TripAdvisor.com (25000 training and 25000 testing reviews for each rating category).

The experiments showed that our rating prediction technique achieve 75-83% in accuracy and can be successfully applied in review filtering systems for suspicious opinions detected.

 © 2014 EffectiveSoft, Ltd.

Practical implications

Intellexer Sentiment Analyzer SDK is the simplest way to understand the voice of your customers

Intellexer Sentiment Analyzer SDK can be effectively used to:

- Capture customer's expectations, attitudes and aversions
- Discover emerging trends
- Analyze feedback on marketing campaigns
- Understand product issues in real-time
- Increase revenue and profits from vendor services
- Gather about competitors' activities
- Understand the emotions of your customers
- Profile customer preferences
- Realize what product or service is better and why

Summarizing benefits

Analyzing the meaning of any opinion in a manner that closely reflects human

Intellexer Sentiment Analyzer SDK delivers:

- Statistical machine learning and semantic technologies to obtain more accurate sentiment analysis results
- Possibility to range documents/opinions by sentiment score
- Faster development time with less costs
- Easy customization to various knowledge domains
- Robustness
- Flexibility
- Possibility to filter non-English text

© 2014 EffectiveSoft, Ltd.