



How an AI-driven recommendation engine can increase Expedia profits by \$5m per annum

In the travel and hospitality markets customers have ever-increasing expectations of a memorable and trouble-free experience. And whilst there is an increasing disillusionment with the reliability of reviews, consumers and providers still want to gain the insight they need to make decisions and take actions.

Applying PrediCX to public reviews on Expedia

Warwick Analytics applied its PrediCX software to publicly available reviews for the travel booking company, Expedia. The aims of the analysis were as follows:

- To improve the profitability of Expedia
- To improve the profitability of Expedia's partners
- To improve the customer experience
- To present better information to customers to improve their decision-making

PrediCX uses Concepts not Keywords

PrediCX is an automated machine learning platform that quickly and accurately generates models for text, using 'human-in-the-loop' technology i.e. it only needs minimum input from a non-data scientist. It took only a few hours to generate meaningful output, no matter how large the dataset, based on concepts instead of keywords and sentiment scoring.

What PrediCX found

Early Warning

The analysis in Figure 1 shows Second Level issues in which over 100 labels were identified. These were grouped by product topic e.g. Car Rental, Hotel, Website, Insurance, Generic and Intents i.e. whether the customer was going to use again, recommend, not use again etc. Some examples of topics include: Website – Freezing; Website – Not Able to Change Dates; Website – Unable to Use Coupons; Hotel – No Air Conditioning; Car Rental - Car Not Matching Order etc.

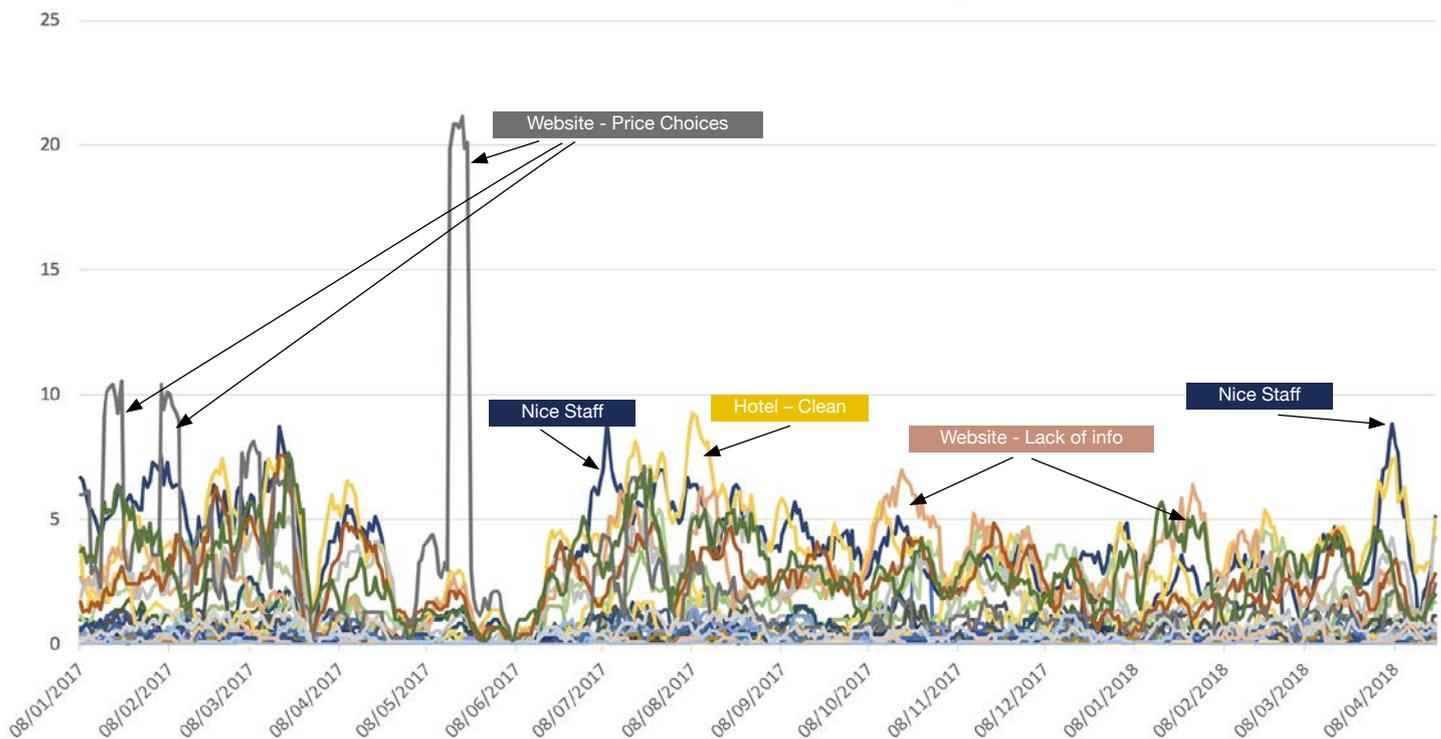
You can also see that there's can also be also a third level of detail accessible. For example, under Website – Price Choices there are the following sub-categories:

- Praise wide range of Expedia's price options
- Praise Expedia's prices that fit their budgets
- Complain about narrow range of Expedia's price choices
- Comment that Expedia have the same prices than other sites

It becomes clear that concept-based insight moves beyond measuring whether someone is 'happy' about a topic.

You can see that there's clearly been a spike in "Price Choices". Whether this is positive or negative sentiment isn't as important as the fact that it is a missed expectation to the customer compared to the brand-promise. Other signals can then be analysed, for example the spikes in "Website - Lack of Info" in Q4 2017 and Q1 2018 appear to relate to email confirmations and hidden charges which were amended. As you can see, a high level of detail can be accessed to get the insight required.

Figure 1.



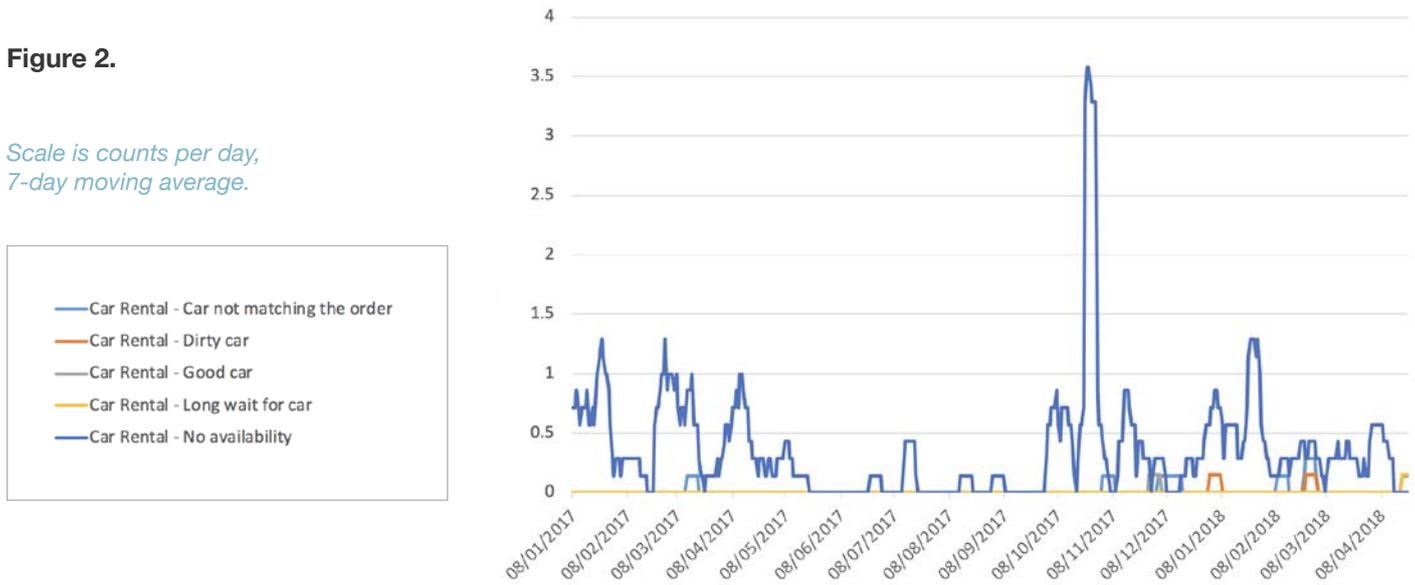
Scale is counts per day, 7-day moving average, not all categories shown.

Early Warnings specifically for Car Hire

Rather than just 'car hire' and 'good' or 'bad' sentiment, the actionable concepts have been identified. You can clearly see that the critical concept is a fleet management challenge, i.e. that there was no availability on particular dates for the car operators in aggregate. The actions are several: predict and pre-empt any seasonal or event-driven shortages, amend and extend the supply base, amend the pricing or communication to align customer expectations. These tags can be matched and passed back to the operators for their own optimisation. It is also possible to make these tags available to customers to give them the choice of which firms to use. This may or may not fit with Expedia's business model but it is a possibility, and can provide trust and transparency.

Figure 2.

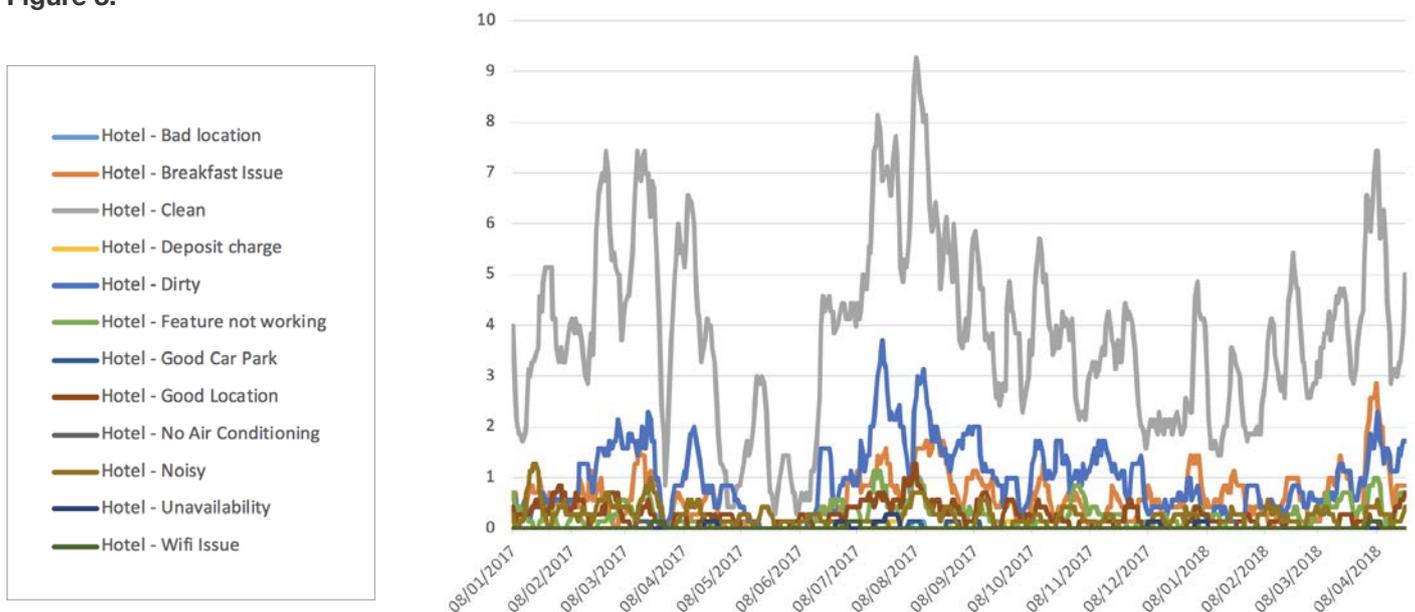
Scale is counts per day, 7-day moving average.



Early Warnings specifically for Hotels

It is a similar picture for hotels, and it is clear that the number one concern is the cleanliness of the hotels. Here the counts are shown at Level 2:

Figure 3.

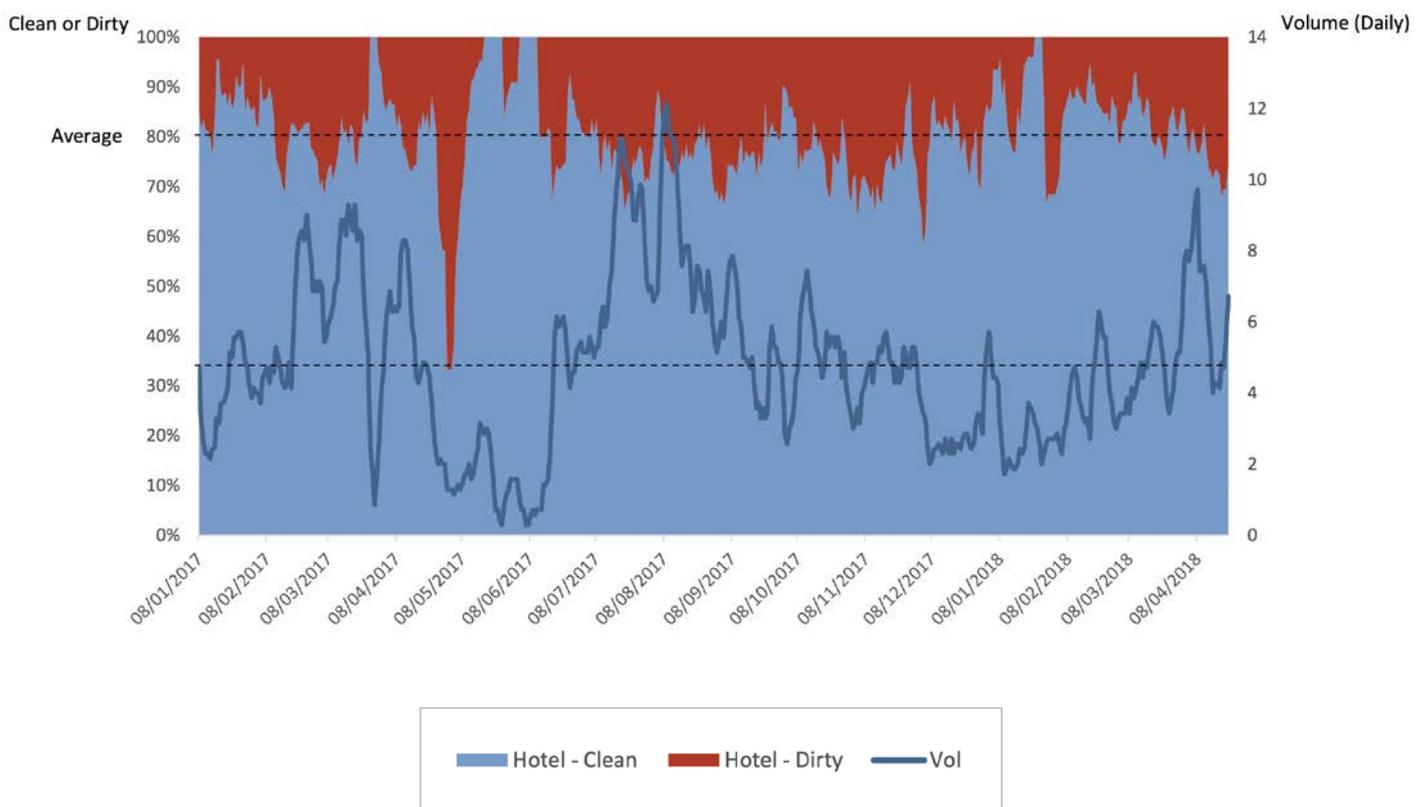


Early Warnings specifically for Hotel Cleanliness

The volume can be adjusted to produce a 'clean-o-gram' in aggregate (shown below) as well as for specific hotels and brands. This is all based on public data, but it can be added to across other private datasets such as surveys and CRM enquiries and complaints.

Figure 4.

Scale is counts per day, 7-day moving average.



Early indicators of customer churn

As referred above, customer intents were also captured such as declared churn as a concept, i.e. customers saying they would never use Expedia again, and developed a dynamic analysis with the declared reason (where there was one) as well as a multivariate analysis where there was no specific declared reason. A simplistic model for the loss of profits from a churning customer was generated from publicly available data (i.e. the accounts) and scaled up from the number of reviews to the number of customers per annum. The output was a 'recommendation engine' showing which factors had the most significant effect on churn, and the impact of a one percent improvement of each of them on Expedia's bottom line.

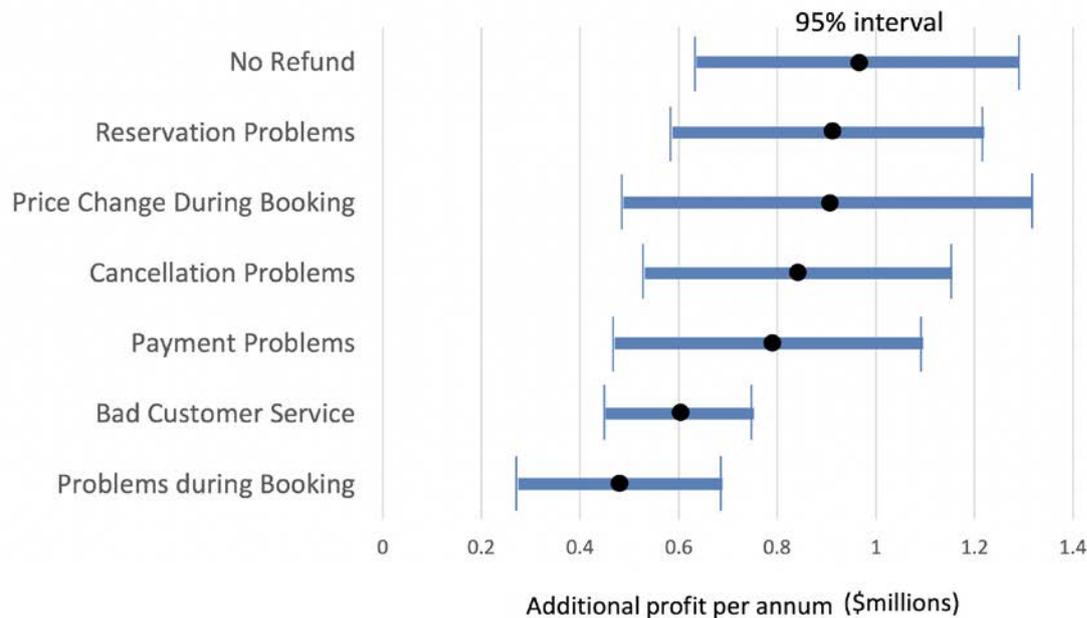
PredicX analysis quickly identified savings of \$millions each year

You can see below that if customers leaving due to 'no refund' was addressed by just 1%, it would add nearly \$1m per annum profit. Similarly, if reservation problems were reduced by just 1%, it would add \$900k per annum profit and so on. Of course, each of these factors would have different ease and cost to change and provide the framework for an executive decision (e.g. it might not be economical to give more people refunds). This analysis can be made more sophisticated by taking internal (private) financial data of customer lifetime value and churn from CRM systems.

Figure 5.

Financial Effect of Improving Top Root Causes of Churn Each by 1%

WA Analysis, Expedia Annual Report



This recommendation engine can also be complemented by a similar model for Expedia's suppliers as well as the lower-level, operational insight from the streams of relevant concepts captured for each part of the customer journey. You'll notice that CSat or NPS hasn't been mentioned until this point. A recommendation engine can also be set up to predict against the factors which drive satisfaction as well as profits, which would also provide insight into how the two are related.

Customer Engagement

PrediCX can be used to pick up the reviews which contain concepts for churn, negative advocacy or the root causes of churn. They can be quickly intercepted by Expedia or its partners to try to recover the customers with an appropriate message or offer, as well as decreasing the negative advocacy on the web. It can also be used for marketing effectiveness, e.g. picking up concepts of where people have used vouchers and the associated experience and loyalty.

Fake Reviews

One of the banes of social media is the growing issue of fake, solicited and gamified reviews. There is no way to stop this entirely, but PrediCX can help to identify fake reviews and remove suspicious or simply glib reviews such as: "everything" [5 stars], or "excellent" [5 stars].

Conclusion

Warwick Analytics was able to generate actionable insight potentially worth many millions of dollars per annum of opportunity for Expedia and its partners, as well as many other use cases to benefit customers too, directly and indirectly.