

Schätzung der touristischen Nachfrage auf Basis von Kundenfeedback in sozialen Medien

Nadine Liedtke,

Forschungsgruppe Data Science, Institut für Digitalen Wandel (IDW)

Einführung

- Statistiken zur touristischen Nachfrage stehen i.d.R. erst mit größer zeitlicher Verzögerung zur Verfügung
- Kundenfeedback im Internet stellt eine alternative Möglichkeit dar, das Reiseaufkommen abzuschätzen
- Mittels Data Mining können Erkenntnisse gewonnen werden, welche sich auf die Planung vorteilhaft auswirken

Methodik

Datenextraktion & Aufbereitung:

- Extraktion und Abspeicherung der relevanten Kundenbewertungen mittels Web-Crawling und RPA
- Text-Preprocessing: Beseitigung von Fehlwerten, Zusammenführen von Datensätzen, Anpassung von Datentypen, etc.
- Aufbereitung von Datensätzen für „normale“ und „Krisen“-Zeiten
 - Zeitraum von 2010 bis 2017
 - Zeitraum von 2010 bis 2020

Komponentenmodell:

- Zerlegung der Zeitreihen (touristische Ankünfte und Kundenbewertungen über obige Zeiträume) in Trend-, Saison- und Restkomponente

Lineare Regression:

- Analyse der Korrelation zwischen touristischen Ankünften und Kundenbewertungen
- Schätzung der touristischen Ankünfte auf Basis von Kundenbewertungen mittels linearer Regression

Validierung der Ergebnisse:

- Gegenüberstellung der Ergebnisse mit alternativen Verfahren (Lineare Trend- und Saisonvorhersage als Basisverfahren)
- Ermittlung des Mehrwerts einer Schätzung touristischer Ankünfte anhand von Kundenbewertungen

Literatur

1. Gebel, A., 2020. Social Media im Tourismusmarketing - Wie Urlaubsanbieter in sozialen Medien Sichtbarkeit und Direktbuchungen steigern. München: SpringerGabler-Verlag.
2. Stotzner, M.-W., 2017. Regressionsanalyse in der empirischen Wirtschafts- und Sozialforschung Band 1 - Eine nichtmathematische Einführung mit SPSS und Stata. Berlin: SpringerGabler-Verlag.
3. Klaus Backhaus, B. E. W. P. R. W., 2018. Multivariate Analysemethoden - Eine anwendungsorientierte Einführung. Berlin: SpringerGabler-Verlag.

Zielsetzung

- Schätzung der Anzahl (vergänger) touristischer Ankünfte auf Basis des Volumens von Kundenfeedback in sozialen Medien
- Bewertung, ob Schätzungsverfahren mittels Kundenrezessionen einen Mehrwert gegenüber klassischen Prognoseverfahren aufweisen

Ergebnis

Anwendungsfall 1: „normale“ Zeiten (2010-2017)

Messwert	Originalreihe	Trendbereinigt	Saisonbereinigt	Restkomponenten
Correlation	0,664	0,596	0,917	0,017
Squared Correlation	0,543	0,154	0,841	0
RMSE	107.036,090	112.478,317	35.805,303	28.284,851

Anwendungsfall 2: Krisenzeiten (2010-2020)

Messwert	Originalreihe	Trendbereinigt	Saisonbereinigt	Restkomponenten
Correlation	0,696	0,688	0,704	0,68
Squared Correlation	0,503	0,275	0,622	0,421
RMSE	152.952,970	139.653,394	177.292,032	105.367,435

Correlation: Korrelation zwischen Ankünften und touristischem Feedback

Squared Correlation: Güte des Modells zur Schätzung der Ankünfte auf Basis des Feedbacks

RMSE: Fehlermaß der Abschätzung der Ankünfte auf Basis des Feedbacks

Schätzung der Kundenankünfte

Messwert	Anwendungsfall eins		Anwendungsfall zwei	
	Schätzung	Prognose	Schätzung	Prognose
Squared Correlation	0,976	0,976	0,805	0,624
RMSE	25.049,036	25.052,541	118.844,999	162.068,008

Verbesserung des RMSE in Prozent

< 0.1%

~ 27%

Schätzung: Schätzung der Kundenankünfte durch Miteinbindung des Kundenfeedbacks

Prognose: Prognose der Ankünfte über Trend- und Saisonkomponente

- Das Modell bietet in normalen Zeiten keine Verbesserung gegenüber klassischen Prognoseverfahren – die Restkomponente ist wertlos und Ankünfte können vollständig aus Trend und Saison abgeleitet werden
- In Krisenzeiten bietet das Modell jedoch eine Verbesserung gegenüber klassischer Prognoseverfahren und kann Besucherzahlen rund 27 % genauer prognostizieren

Kontakt

Prof. Dr. Wolfram Höpken
Leiter IDW / Forschungsgruppe Data Science
+49 751 501 9764
wolfram.hoepken@rwu.de