

Query Abandonment Prediction with Recurrent Neural Models of Mouse Cursor Movements

Lukas Brückner, Aalto University

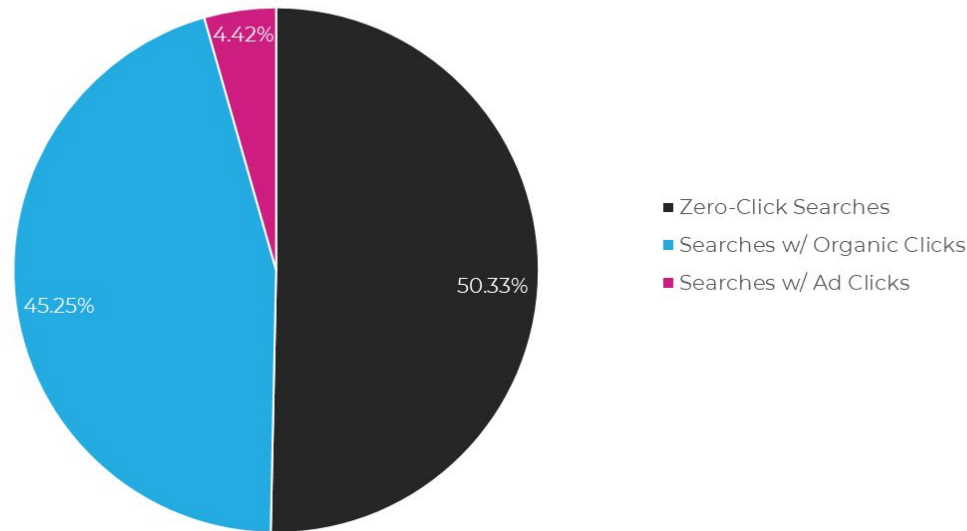
Ioannis Arapakis, Telefonica Research

Luis A. Leiva, Aalto University

Less than half of Google Searches result in a click

Paid, Organic, & Zero-Click Searches in Google (June 2019)

data from 40M+ browser-based searches on millions of desktop & mobile devices in the United States




jumpshot

SparkToro

Source: <https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/>

“Good” query abandonment

 ✕ 🔊 🔍 Sign in


[All](#) [Images](#) [News](#) [Videos](#) [Shopping](#) [More](#) [Settings](#) [Tools](#)

About 62.100.000 results (0,70 seconds)


Brad Pitt / Age


56 years


December 18, 1963



People also search for

 **Tom Cruise**
58 years

 **Jennifer Aniston**
51 years

 **George Clooney**
59 years

Feedback

Brad Pitt

American actor

William Bradley Pitt is an American actor and film producer. He has received multiple awards, including two Golden Globe Awards and an Academy Award for his acting, in addition to another Academy Award and a Primetime Emmy Award as producer under his production company, Plan B Entertainment. [Wikipedia](#)

Born: December 18, 1963 (age 56 years), [Shawnee, Oklahoma, United States](#)

Height: 1.8 m

Full name: William Bradley Pitt

Children: [Shiloh Jolie-Pitt](#), [Maddox Chivan Jolie-Pitt](#), [MORE](#)

Previous work relied on engineered features

Leaving So Soon? Understanding and Predicting Web Search Abandonment Rationales

Abdijogiri Diriyel, Ryan W. White², Ganga Raghav², and Susan T. Dumais²

6.1 Feature Generation

Our predictions are made retrospectively, after abandonment is observed. Around

2,000 features

ABSTRACT
Users of search engines
high frequency of
Web search engines

A. Diriyel, R. White et al. 2012. Leaving So Soon? Understanding and Predicting Web Search Abandonment Rationales

Predicting User Engagement with Direct Displays Using Mouse Cursor Information

Table 2: Features used in the classification task to predict user engagement.

Base features	Meta-features	Aggregate functions*
Viewport (width, height)	# Moves (towards, away) KM	x_{min}, x_{max}
Cursor positions and timestamps	# Moves (towards, away) KM within dist. d	$\sum \mu, \bar{x}$
Unique cursor positions	# Clicks (inside, outside) KM	$\sigma^2, \sigma_{x1}, SST$
Normalised viewport positions	Time to first click on KM	\sum intra-distance
Unique normalised viewport pos.	# Preceding clicks to KM	Shannon entropy
Subsequent points' distance	# Hovers over KM	Permutation entropy
Subsequent points' duration	# Hovers over KM	Permutation entropy
Cursor distance from KM	# Preceding hovers	μ_{10} most p
Cursor speed	Time to first hover (KM, other elements)	Multivariate KL
Cursor normalised speed	Time hovering (KM, other elements)	Earth mover's distance
Cursor acceleration		

638 features

I. Arapakis and L. A. Leiva. 2016. Predicting User Engagement with Direct Displays Using Mouse Cursor Information

Does That Mean You're Happy?

RNN-based Modeling of User Interaction Sequences to Detect Good Abandonment

Kyle Williams and I. Zitouni

3.2 Interactions

In this study, we consider the following user interactions that could potentially occur on an abandoned SERP.

Short Pause (SP) The user paused on the SERP for $s \leq 5$ seconds.

Medium Pause (MP) The user paused on the SERP for $5 < s \leq 15$ seconds.

Long Pause (LP) The user paused on the SERP for $15 < s \leq 30$ seconds.

Very Long Pause (VLP) The user paused on the SERP for $s > 30$ seconds.

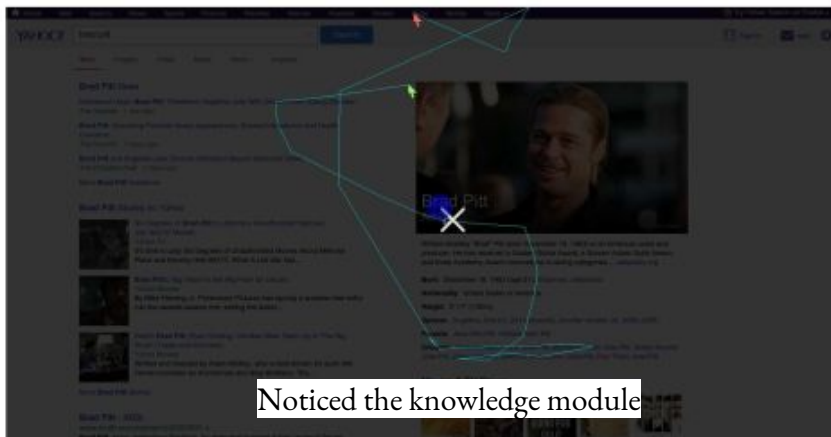
sequence of 10 features

ABSTRACT

Queries for which there are no subsequent queries. Differentiating between good and bad abandonment has become an important problem for search engines. We require users to click. Existing work on the SERP and detailed investigation of how sequence of user interactions between good and bad abandonment patterns on a label

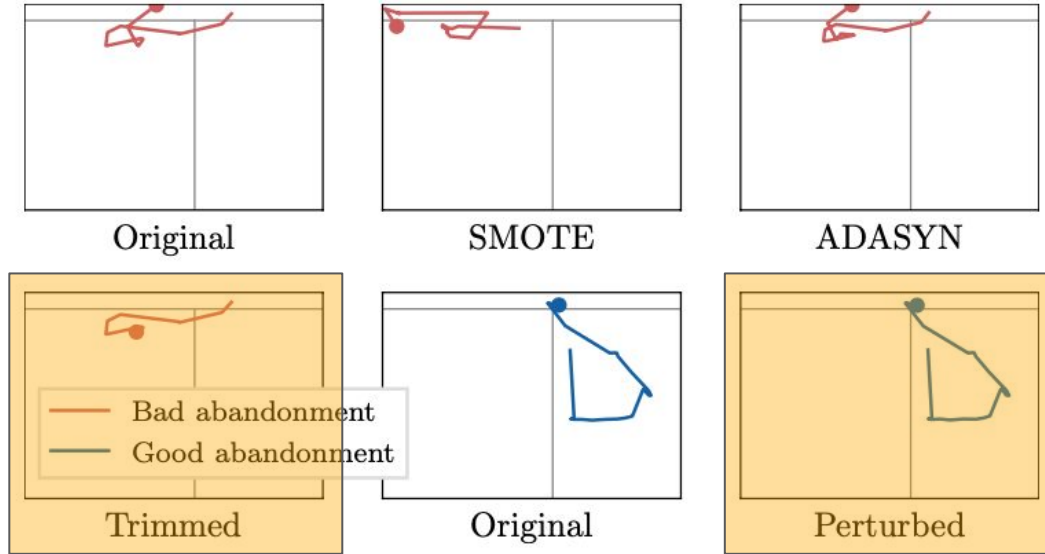
K. Williams and I. Zitouni. 2017. Does That Mean You're Happy? RNN-Based Modeling of User Interaction Sequences to Detect Good Abandonment

Mouse Movements allow to understand user behavior*



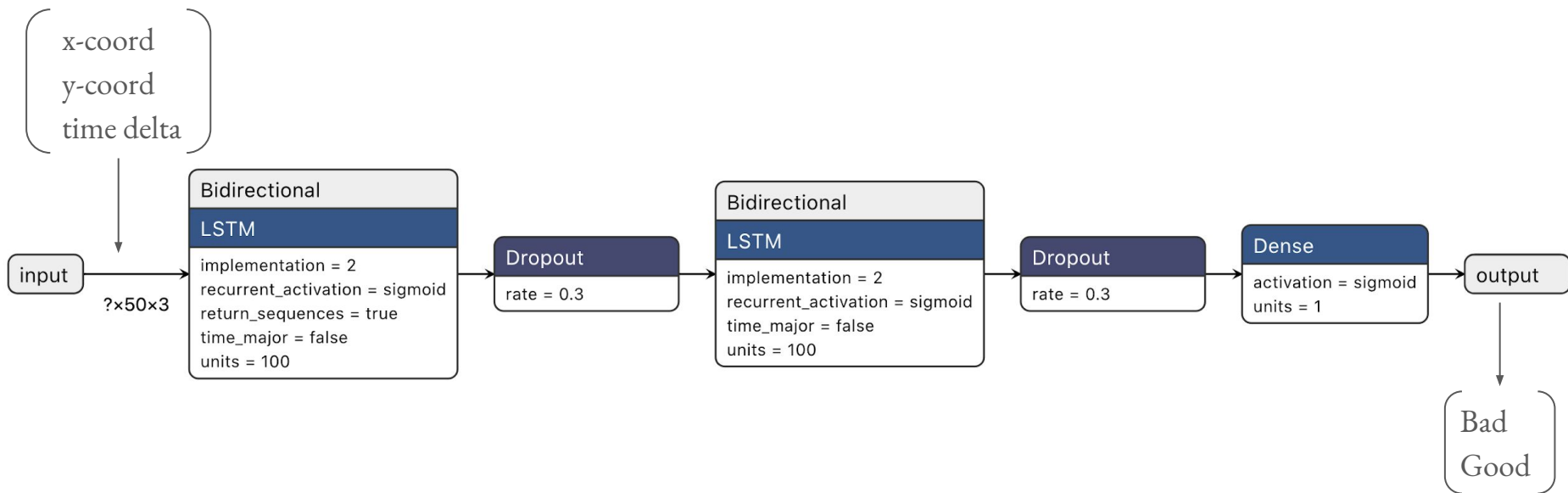
→ No need for expensive, handcrafted features

Domain-specific data augmentation helps prediction



Binary Classification with BiLSTM Model

Last 50 steps



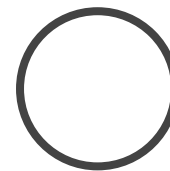
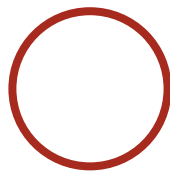
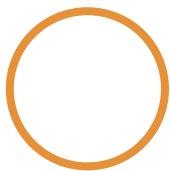
Highest ROC AUC with custom augmentation

Input data	Time	Augmentation	Adj. Precision	Adj. Recall	F-measure	ROC AUC
<i>All abandoned queries are considered bad abandonments</i>			0.08 [0.06, 0.11]	0.28 [0.24, 0.32]	0.12 [0.10, 0.15]	0.50 [0.50, 0.50]
RF using 10-dim feat vectors		ADASYN	0.67 [0.58, 0.76]	0.64 [0.54, 0.73]	0.64 [0.54, 0.73]	0.60 [0.50, 0.69]
XGB using 10-dim feat vectors		ADASYN	0.70 [0.60, 0.78]	0.65 [0.55, 0.74]	0.65 [0.55, 0.74]	0.61 [0.51, 0.70]
Standardized coords	no	none	0.52 [0.48, 0.57]	0.72 [0.68, 0.76]	0.60 [0.56, 0.65]	0.50 [0.46, 0.55]
Raw coords	yes	rand. undersample	0.68 [0.64, 0.72]	0.59 [0.54, 0.63]	0.59 [0.54, 0.63]	0.59 [0.54, 0.63]
Standardized coords	yes	rand. oversample	0.67 [0.62, 0.71]	0.63 [0.58, 0.67]	0.62 [0.57, 0.66]	0.59 [0.54, 0.63]
Speed only	implied	SMOTE	0.67 [0.63, 0.71]	0.63 [0.58, 0.67]	0.63 [0.58, 0.67]	0.59 [0.55, 0.63]
Speed + distance to KM	implied	SMOTE	0.70 [0.65, 0.73]	0.65 [0.61, 0.69]	0.65 [0.61, 0.69]	0.61 [0.57, 0.65]
Raw coords	yes	SMOTE	0.69 [0.65, 0.73]	0.63 [0.59, 0.67]	0.63 [0.59, 0.68]	0.61 [0.57, 0.65]
Standardized coords	yes	ADASYN	0.68 [0.64, 0.72]	0.64 [0.60, 0.68]	0.64 [0.59, 0.68]	0.61 [0.56, 0.65]
Standardized coords	yes	distortion or trimming	0.72 [0.68, 0.76]	0.65 [0.61, 0.69]	0.65 [0.61, 0.69]	0.63 [0.59, 0.68]

Table 1: Experiment results. Top rows are baseline conditions. We report the best combination of {Coords, Time, Resampling, Augmentation} techniques tested (36 in total). 95% conf. intervals according to the Wilson method for binomial distributions.

→ Slightly improving on baselines based on engineered features of previous work

Discussion



Targeted Models

Commercial providers might want to ensure only bad abandonments are investigated

Limitation

Data set with small sample size

Practical

Mouse movements for abandoned queries complement click-based metrics for SERPs

Thank you for your attention

This presentation was supported by the **SIGIR Student Travel Grant**

Slides: <https://github.com/luksurious/abandonment-rnn/presentation-cikm.pdf>

Code: <https://github.com/luksurious/abandonment-rnn>

Talk to us on Twitter: [@luksurious](#) [@iarapakis](#) [@luileito](#)

Contact: lukas@lukas-brueckner.de