Google Keyword Cost-Per-Click Analyzer

Mohammad Sabir Bepari, Shivanand Ekbote, Saifali Kalkeri,
Yash Purohit, Shivakumar Hiremath and Priyadarshini Patil

February 8, 2024

Google Keyword Cost-Per-Click Analyzer

Mohammad Sabir Bepari[1], Shivanand Ekbote[1], Saifali Kalkeri[1], Yash Purohit[1], Shivakumar Hiremath[1], and Ms. Priyadarshini Patil[1]

School of Computer Science and Engineering, KLE Technological University, Hubballi, India

Abstract. This leading utilization and achieving of online advertising campaigns using advanced keyword CPC (cost-per-click) analysis techniques. This research focuses on predictive modeling, bid optimization, and anomaly detection to improve advertisers' decision-making process. Understanding the factors that influence CPC allows advertisers to strategically allocate budget, adjust ad copy, and refine targeting techniques. This research highlights the importance of continuous learning that allows campaigns to adapt to evolving market dynamics. Overall, the results demonstrate how a nuanced approach to CPC analysis can improve the efficiency of online advertising, leading to improved performance and resource utilization.

1 Introduction

In the evolving world of online advertising, collaboration between publishers and advertisers plays a critical role in achieving success. Within this complex ecosystem, the pricing of Google keywords through Cost Per Click (CPC) holds significant weight, directly affecting both publishers' revenue and advertisers' budget optimization. Accurately anticipating and comprehending CPC patterns is vital for making prudent choices, which is the focus of this project. In this study, we explore the dynamic relationship between Search Engine Optimization (SEO) and the creation of a machine learning model designed to forecast Google keyword CPC. SEO is a crucial element for improving online presence, drawing in natural traffic, and optimizing websites to rank pertinent keywords.
Simultaneously, forecasting CPC values grants publishers the ability to strategically position ads to maximize revenue, while giving advertisers the opportunity to optimize their spending by targeting cost-efficient keyword In this lively discussion, we take an in-depth look at the vital role of search engine optimization (SEO) in improving visibility and attracting visitors. We examine how search engines ¬s process and rank webpages to uncover organic traffic sources. Implementing SEO is the fist practice that allows sites to rank higher in the results. Understanding search engine algorithms and providing quality content offer immense value in the digital age. The SEO websites are found by relevant audiences. As we move forward, our attention shifts to developing a robust machine-learning model capable of accurately forecasting the (CPC) for Google keywords. The ultimate objective is to reliably predict future CPCs, arming publishers, and

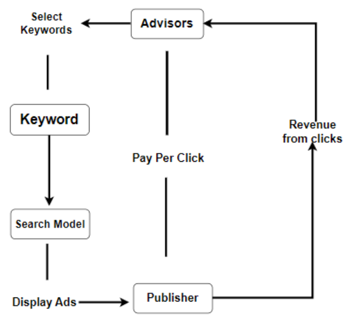

Fig. 1. Flow diagram of CPC

advertisers with data-driven insights to make smart decisions. The model forecasts the CPCs down the line by leaving the historical CPC data and relevant factors. These empower key players to navigate the fast-paced, competitive digital advertising arena with greater confidence. By combining past learning with present-day inputs, actionable CPC predictions can be used to guide stakeholders operating in this dynamic landscape. We sometimes discuss how crucial the SEO is for Higher Education Institutions (HEIs). Web visibility and technological innovation significantly impact their achievement and reputation. A research project that aims to examine SEO review outcomes on HEI sites, contrast them locally, and suggest better plans is explained. Combining SEO knowledge and AI forecasts, this document intends to offer an all-e-encompassing view of how using data can boost results for both content creators and advertisers in the complex world of Internet ads

2 Literature survey

Various methods and algorithms have been studied in the context of CPC prediction. Previous studies have used rule-based systems and various machine learning models, but challenges remain in achieving high accuracy and adaptability to dynamic advertising environments. Our study builds on these efforts and introduces the use of random forest models for CPC prediction. This approach aims to leverage the model's ability to capture complex relationships in the data and provide a more nuanced understanding of the factors that influence CPC.
1. "Optimize increase in cost per click using reinforcement learning without exploration" This article explores the application of reinforcement learning to optimize enhanced cost-per-click (ECPC) for online advertising without explicit research. Reinforcement learning, a machine learning paradigm inspired by behavioral psychology, is used to make decisions in dynamic environments. The

main contribution of this research lies in the development of a reinforcement learning model that optimizes his ECPC without the need for explicit exploration, potentially leading to more efficient and targeted advertising campaigns. This study provides valuable insights into the interface between reinforcement learning and online advertising optimization. [1]
2. “Online advertising pricing: cost-per-click-through and cost-per-action.” In this article, the authors undertake a comparative analysis of his two main pricing models in online advertising: cost per click (CPC) and cost per action (CPA). This study explores the pros and cons of each pricing model and reveals the factors that influence the choice between CPC and CPA. By investigating pricing dynamics, this paper contributes to a deeper understanding of how advertisers can align their advertising objectives with the most appropriate pricing strategy. This knowledge is critical for marketers who want to make informed decisions about online advertising spending.[2]
3. “Third Party Adwords Management in SEM: Optimization Models and the Potential of Twitter” This document focuses on AdWords management for third-party search engine marketing (SEM) and presents an optimization model. This study not only outlines a model for improving AdWords campaigns, but also investigates the possibility of incorporating Twitter into your optimization strategy. This paper provides an advertiser with a systematic approach to effectively manage her AdWords campaigns by proposing an optimization framework. Exploring the potential of Twitter adds an innovative dimension and suggests a holistic approach to online advertising that goes beyond traditional platforms.[3]
Together, these three of his articles contribute to the existing body of knowledge about online advertising and his SEM. These provide insight into optimization strategies, pricing models, integration of new platforms such as Twitter, and provide a comprehensive background for the development and understanding of tools such as the 'Google Keyword CPC Analyzer'.
Various methods and algorithms have been studied in relation to CPC prediction. Previous studies have used rule-based systems and various machine learning models, but challenges remain in achieving high accuracy and adaptability to dynamic advertising environments. Our study builds on these efforts and introduces the use of random forest models for CPC prediction. This approach aims to leverage the model's ability to capture complex relationships in the data and provide a more nuanced understanding of the factors that influence CPC.

3 Proposed methodology

This project examines various data regarding different keywords used in online advertising. Our goal is to find out which keywords have the potential to generate more revenue for advertisers when users click on their ads. By examining the data, we can recommend to investors which keywords are more profitable, allowing them to make better decisions about where to invest their advertising dollars for maximum return. We will assist you in this way. It's like becoming a detective in the world of online advertising and finding hidden gems that will make advertisers more successful. In this project we are using the machine learning model as Random forest model.

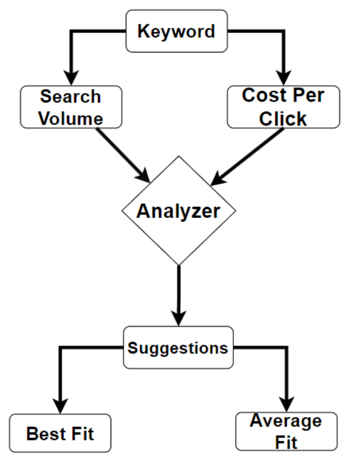

Fig. 2. Proposed methodology

Backbone

3.1 Gathering Dataset

The initial phase involves searching in google for a proper dataset which consist of the relavant table of data which can support our dataset and can help in the proper training of the model which can give us the proper resulted expected. we have obtained the dataset from the google website in adsense we have gathered the datase from the website answer the public where we can filter the keyword according to the needs of our search. We are able to download the filted keyword which will be in the form of csv where it contains the frequent or most searched keywords The dataset size can be exponential increased as we can add more no of keywords in future.

3.2 Preprocessing

In this study focused on keyword Cost per click analysis of online advertising, of the dataset is a key aspect to ensure accurate and insightful results. The dataset, consisting of keywords and associated Cost per click values, undergoes a careful cleaning procedure. Missing data is identified and imputed or removed to ensure the completeness of the data set. Standardization uses keyword normalization techniques such as lowercase conversion and special character removal. Categorical variables, such as keyword categories, are numerically coded to facilitate their integration into analyses. Scaling of Cost per click values is implemented to compensate for possible differences using methods such as min-max scaling and standardization to ensure consistency. Outliers in Cost per click values are inspected and adjusted or removed as appropriate. When applicable, time-based features are treated with care and irrelevant features are cleaned up to optimize the data set. Exploratory data analysis and visualization techniques are used to gain insights and identify patterns that form the basis for further decisions. Feature engineering is the introduction of new variables that improve a model's predictive capabilities. This comprehensive framework ensures a clean and standardized dataset, laying the foundation for robust analysis and informed decision-making in the field of online advertising keyword Cost per click analysis.

3.3 Model

The Random Forest algorithm is an ensemble learning method that constructs a multitude of decision trees. Each tree is trained on a different subset of the data, and the final prediction is an aggregation of individual tree predictions. This ensemble approach enhances the model's robustness and ability to generalize to unseen data. Random Forests, the concept of bagging, or bootstrap aggregation, is integral. This technique initiates with the utilization of randomly selected data, which is then organized into sets termed bootstrap samples through a process known as bootstrapping. Subsequently, independent models are trained on these bootstrap samples, generating diverse outcomes or aggregations. In the final stage, the collective results are combined, and the output is determined based on a majority vote. This entire process, involving the training of independent models and the amalgamation of results through a majority vote, is facilitated by an ensemble classifier, commonly known as bagging.

Backbone The architecture and structure of each decision tree in the ensemble is referred to as the Random Forest model's backbone. In this unique situation, every choice tree is viewed as a base student or a feeble student. The spine is intended to catch complex connections and examples inside the information.

Structure of a Tree: Every choice tree in the Arbitrary Timberland is developed through recursive double parts. The parts depend on include values, and the tree not entirely set in stone by choosing the highlights that give the best partition of information at every hub.

Group Approach: The force of the Arbitrary Woodland lies in its troupe approach. Via preparing different choice trees freely and accumulating their expectations, the model turns out to be more hearty, decreasing overfitting and upgrading speculation to concealed information.

Detecting Head The part of the Random Forest model that is in charge of combining individual decision tree predictions into a single prediction is called the detecting head. Every choice tree doles out a class or a relapse worth to an important piece of information, and the identifying head totals these singular forecasts.

Grouping Undertaking: In a grouping task, the recognizing head frequently utilizes a greater part casting a ballot component, where the class with the most votes across all choice trees turns into the last anticipated class.

Relapse Assignment: In a relapse task, the distinguishing head normally midpoints the relapse values relegated by individual trees to get the last expectation.

Mosaic Data Augmentation A technique called mosaic data augmentation is employed, particularly in computer vision tasks, to improve the Random Forest model's robustness and conceptual capacities. Building Mosaic Images Using a variety of training images, mosaic images are created using mosaic data augmentation. Patches or sub-pictures from various original photographs are stitched together to achieve this. The Random Forest model is trained using the mosaic images as well. spatial relationships Through the use of spatial relationships between objects and backdrops from colorful photographs, Mosaic Data Augmentation aids in the model's learning to handle a variety of settings and object placements, increasing its scriptability. A Greater Degree of Diversity By adding mosaics to the training set, the variability of the data is increased, which hinders the model from seeing patterns in individual images and honing its ability to generalize

4 Implementation

Each tree in the Random Forest is trained on a bootstrapped sample of the original dataset. This involves aimlessly slice with relief from the training set, creating a different set of training data for each tree. At each split in a decision tree, only a subset of features is considered. This introduces randomness and decorrelates the trees, precluding overfitting to specific features. Train each decision tree using its separate bootstrapped sample. The tree is grown by recursively unyoking bumps grounded on the named features until a stopping criterion is met(e.g., maximum depth or minimum samples per splint). Combine prognostications from individual decision trees to make the final vaticination. For bracket, this frequently involves a maturity vote, and for retrogression, it's generally the normal of individual prognostications.

5 Metrics

1. Impression: Every time an ad appears on Google's search result page, it is counted as an impression. Explanation: It tracks how often people see an ad when searching Google.

2.Click: When someone clicks on an ad (like the blue headline of a text ad), Google counts it as a click. Explanation: This shows how many times people interact with an ad by clicking on it.

3.Conversion: A conversion occurs when someone clicks on an ad and takes an action valuable to your business, such as making a purchase or calling you. Explanation: It measures the meaningful actions that users take after clicking on an ad.

4. Click-through rate (CTR): The percentage of clicks that your ad receives out of the total times it is shown (clicks/impressions). Explanation: A high CTR indicates that users find their ads to be helpful and relevant.

5. Conversion rate (CR): The average number of conversions per ad click was expressed as a percentage (conversions/clicks). Explanation: This tells you how effective your advertisements are in turning clicks into valuable actions.

6. Maximum CPC (bid): The highest amount you are willing to pay for a click on your ad. Explanation: This helps control advertising costs by setting a limit on how much you are willing to pay for each click.

7. Average CPC: The average cost you have been charged for a click on your ad (total cost of clicks/total number of clicks). Explanation: It shows the average amount you are paying for people to click on your ad.

8. Quality Score: An estimate of how relevant and useful Google thinks an ad and landing page is for a specific keyword.

6 Results

SL.NO	SUGGESTIONS	SEARCH VOLUME	KEYWORDS	COST PER CLICK
1.	where's google drive	140	google	20.42
2.	which google pixel is best	170	google	0.04
3.	how google search works	390	google	10.05
4.	what google analytics do	30	google	13.25
5.	google translate	37200000	google	0.19
6.	google earth	1000000	google	0.02
7.	google youtube ads	5400	google	443.6
8.	google youtude	22200	google	0.03
9.	google your favourite cricketer	1000	google	0
10.	google yah batao	8100	google	0

Fig. 3. Results of the analysis

To evaluate the performance of our Keyword CPC Analyzer, we conducted experiments using real-world CPC campaign data. The dataset included information on keywords, historical performance, and competitive bid landscapes. We compared the results of our Random Forest model with traditional CPC analysis methods.

Google Keyword CPC Analyzer is designed to provide advertisers and publishers with valuable insight into keyword cost-per-click (CPC) trends, allowing them to make informed decisions and optimize their advertising strategies. It

has been. This section describes the expected results and possible impacts of using this tool.
1. Improved CPC prediction: The main result of "Google Keyword CPC Analyzer" is that it allows you to predict her CPC value for a particular keyword more accurately and reliably. Machine learning models trained on historical data and market trends identify patterns and fluctuations in CPC, leading to improved predictions. This improved predictive capability allows advertisers to allocate budget more effectively and bid strategically on keywords.
2. Optimized advertising costs: Accurate CPC predictions allow advertisers to optimize their ad spend by strategically adjusting bids based on keyword volatility. By identifying trends and patterns, this tool helps you make data-driven decisions about when to bid higher or lower for specific keywords. This optimization contributes to more efficient resource allocation and allows advertisers to get the most out of their advertising budget.
3. Maximize sales: The effects of "Google Keyword CPC Analyzer" are not limited to cost optimization. Advertisers can strategically bid on keywords that are likely to generate higher returns, allowing them to target their advertising efforts to keywords that are more likely to generate higher returns. This revenue-driven approach contributes to the overall success of your advertising campaign and maximizes the effectiveness of your advertising dollars spent.
4. Improved decision making: The tool's user-friendly interface provides quick and easy access to CPC predictions, allowing advertisers and publishers to make informed and timely decisions. The results provided by this tool serve as a valuable resource in your decision-making process and help you choose keywords that fit your campaign goals and target audience.
5. Continuous improvement: Google Keyword CPC Analyzer integrates a user feedback loop to enable continuous improvement. As users interact with the tool and provide feedback, developers can iterate on the model to improve its accuracy and usability over time. This commitment to continuous improvement keeps our tools relevant and effective in a dynamic online advertising environment.
Achieving 100 accuracy in a Random Forest model may indicate potential overfitting, where the model perfectly memorizes the training data but struggles with unseen data. To address this, regularization techniques, such as adjusting tree depth and sample requirements, are implemented. Regularization helps control model complexity, enhancing generalization and preventing overfitting, ensuring accurate predictions on new, unseen data. After using regularization the search volume accuracy is 82.67 and the cost-per-click(CPC) is 97.39.

7 Conclusion

In summary, 'Google Keyword CPC Analyzer' is a powerful solution for advertisers and publishers who want to optimize their online advertising campaigns. The tool provides accurate cost-per-click (CPC) predictions through advanced machine learning models, allowing users to strategically allocate budget, maximize revenue, and make informed decisions. Key insights and contributions of this tool include improved CPC predictions, ad spend optimization, improved decision making, and continuous improvement through user feedback. The effectiveness of this tool extends beyond cost optimization to facilitating a user-friendly interface that advertisers can interact with and gain actionable insights. Positive outcomes include improved decision-making, maximizing revenue, and focusing on user satisfaction, leading to higher adoption rates. Future research and development areas lie in the continued development of predictive models. Integrating advanced algorithms and techniques, such as ensemble methods and deep learning architectures, further improves the accuracy of the tool and extends its predictive capabilities. Explore additional advertising metrics beyond CPC, including: Measurements such as click-through rates and conversion rates provide a more comprehensive view of campaign performance, aligning with the industry's need for holistic insights. Additionally, future research can address the integration of new technologies such as natural language processing and voice-activated interfaces to improve user experience and accessibility. Privacy and security remain important, so continued efforts should focus on putting in place strong measures to protect user data and ensure compliance with evolving regulations. In summary, 'Google Keyword CPC Analyzer' is a valuable tool for the evolving online advertising landscape. Future research should aim to expand the boundaries of predictive analytics, incorporate new technologies, and continue to prioritize user-centered design and privacy to meet the dynamic needs of advertisers and publishers in the digital marketing space.

8 Reference

1. Li, Sinan Yuan, Chun Zhu, Xin. (2021). Optimizing Enhanced Cost Per Click via Reinforcement Learning Without Exploration.
2. Yu Hu,Jiwoong Shin, Zhulei Pricing of Online Advertising: Cost-per-Click-through vs.Cost-per-Action Tang
3. Wang, Dong Li, Zhenyu Xie, Gaogang Kaafar, Dali Salamatian, Kave. (2016). Adwords management for third-parties in SEM: An optimisation model and the potential of Twitter
4. Maximum CPC bid, Manual CPC bidding - sup-port.google.com
5. S. Athey and G. Ellison, "Position Auctions with Consumer Search" in , Working Paper, Harvard University, 2008.
6. S. Balachander and K. Kannan, "Pricing of advertisements on the Internet" in Working Paper, Purdue University, 2006.
7. D. Hallerman, "Online ad pricing", Entrepreneu Magazine (March 25), 2002.
8. M. Helft, "Google tests an ad idea: Pay only for results", New York Times, Mar 2007.