Recommendation System for Research Grants

Rakesh Tholiya, Kuldeep Rathore, Mukul Upadhyay, Sachin Choudhary

*Army Institute of Technology, Pune*

**Abstract**

This research paper proposes the design and implementation of a recom- mendation system for researchers to simplify the process of finding research grants. Despite the positive impact that researchers and scientists have on the world through their discoveries and innovations, securing funding for their projects remains a challenging and time-consuming task. This paper aims to address this issue by introducing a recommendation system that leverages the architecture of the system, detailed design, project planning, and require- ments to suggest grants that researchers may find interesting based on their platform activities. The proposed system will be beneficial for researchers, students, and scientists seeking grants, as it will provide an efficient and user- friendly way of finding grants that fit their interests and project requirements.

# Introduction

A recommendation system for research grants is an intelligent system that uses data mining and machine learning algorithms to suggest potential research grant opportunities to researchers based on their interests, skills, and qualifications. The system collects and analyzes data from various sources, including past research publications, conference attendance, and funding history to make recommenda- tions that match the researcher’s profile and the grant’s requirements.

The primary goal of a recommendation system for research grants is to improve the efficiency and effectiveness of the grant application process by providing re- searchers with personalized recommendations, increasing the likelihood of suc- cessful grant applications. The system can also help funding agencies by identify- ing the most promising research projects and facilitating the review and selection process.

To develop an effective recommendation system for research grants, several fac- tors must be considered, including data quality, algorithm selection, and evalu- ation metrics. The system’s data should be accurate, relevant, and up-to-date to ensure the recommendations are tailored to the researcher’s profile and the grant’s requirements. The algorithm selection should also be carefully considered, as dif- ferent algorithms have varying strengths and weaknesses and may perform differ- ently on different types of data.

To evaluate the effectiveness of the recommendation system, various metrics can be used, such as precision, recall, and F1 score. Precision measures the proportion of recommended grants that are actually relevant to the researcher, while recall measures the proportion of relevant grants that are recommended. The F1 score

is a combination of precision and recall that provides an overall measure of the system’s performance.

# Literature Review

## Personality-Aware Product Recommendation System Based on User Interests Mining and Meta path Discovery

* + - Every modern social network or online store must include a recommenda- tion system. The product recommendation system, a classic illustration of the legacy recommendation systems, has two significant drawbacks: sug- gestion repetition and unpredictability about new goods (cold start). These restrictions exist because the older recommendation algorithms only utilise the user’s existing purchasing history to suggest new products.
    - It may be possible to reduce the cold start and eliminate redundant recom- mendations by including the user’s social attributes, such as personality traits and areas of interest. As a result, in this paper, we offer Meta-Interest, a personality-aware product recommendation system based on user interest mining and meta path discovery. Meta-Interest predicts the user’s hobbies and the objects connected with these interests even if the user’s history lacks these or comparable items. This is accomplished by assessing the user’s sub- ject interests and, ultimately, proposing products related to the user’s inter- est.
    - This proposed solution is personality-aware in two ways: it uses the person’s character features to forecast the user’s subjects of interest and it matches the user’s character facets with the linked objects. The suggested system was compared to contemporary recommendation approaches, such as deep- learningbased and meeting recommender. According to experimental find- ings, the suggested strategy can improve the recommendation system’s mem- ory and precision, particularly in cold-start conditions.

## Multiple Time Series Perceptive Network for UserTag Suggestion in Online Innovation Community

* + - In the online innovation community, the user tag suggestion approach, which aims to learn users’ preferences over knowledge items from their history ac- tions, plays a vital part in creating tailored recommendations.
    - Most existing user tagged solutions only employ a single sort of behaviour to forecast a single label for users, resulting in poor generalisation of user profile.In this study, we present a multiple time series perceptive network (MTSPN) for user labelling activities in the online innovation community. The MTSPN specifically looks at a range of user behaviours for the teamwork view.
    - To estimate numerous items for users, the proposed MTSPN technique incor-

porates a cross-classifying module as well as multi-scale incremental features that are drawn from various sequential behaviours. Using only a realworld dataset compiled from the ”Thingiverse” community, our encouraging exper- imental results confirm the superiority of our MTSPN model over a variety of alternative user tagging techniques.

## Location Aware Keyword Query Suggestion Based on Document Proximity

* + - Users can find relevant information using keyword suggestions in web searches without needing to know how to articulate their requests explicitly. The ad- dresses of the users as well as the query results are not taken into account by the currently used keyword suggestion algorithms, i.e., the spatial closeness of a user to the obtained results also isn’t taken into account while making the recommendation.
    - It is well known that in many applications (such as location-based services), the quality of search results relates to their geographical location to the query issuer. In this study, we develop a framework for location-aware keyword query suggestions. Suggest a weighted search term graph that accounts for both the spatial separation between the user location and the returned doc- uments as well as the semantic significance between keyword queries.
    - To choose the keyword queries with the best scores as suggestions, the graph is explored in a random-walk-with-restart method. We provide a partition- based strategy that improves the basic method by up to a magnitude of one in order to make the framework scalable. Real data is used to assess the efficacy of our system and the efficiency of the algorithms.

## Developing a Meta-Suggestion Engine for Search Queries

* + - Typically, search engines offer query recommendations to help users with their searches. It’s crucial to use query suggestions to make users’ search experiences better. The majority of query suggestions, however, are depen- dent on the user’s search history and may be influenced by rarely used search terms. Depending on the user’s search, query recommendations may work well in domestic search engines but not in international search engines.
    - On the other hand, it can be weak in local engines and strong in global en- gines. Furthermore, it is challenging to build log-based query suggestions outside of a big search engine because they necessitate a lot of search records. Some search results do not offer query recommendations, which makes it challenging for users to conduct searches. The user’s search experience is negatively impacted by these query suggestion flaws. In this work, we cre- ate a meta-suggestion, a fresh approach to query recommendation. Meta- suggestions retrieve potential questions and suggestions from those other search engines, just like meta searches do.
    - By re-ranking the collected candidate query results, meta-suggestions pro- duce suggestions. We create a browser extension called the meta-suggestion

engine (MSE) that produces meta-suggestions. It doesn’t need a search log and can propose queries for any web page. Our meta-suggestions performed 17on normalised discounted cumulative gain (NDCG) and 31when compared to well-known search engines like Google. It is anticipated that user queries will significantly advance if more factors, including such user preferences, are found through continuous research. If aspects like user preference are looked at in further research, an improved user search experience may be achievable.

## BTR: A Feature-Based Bayesian Task Recommendation Scheme for Crowdsourcing System:

* + - The crowdsourced system is a platform for distributed problem-solving where jobs are posted as an open call and dispersed to a crowd (i.e., crowd workers). Huge-scale crowdsourcing systems typically contain many microtasks, and the time a crowd worker must spend finding the right task may be equal to the price of completing the task. Task suggestion is therefore required.
    - Current work ignores the crowdsourcing system’s dynamics, or the fact that new jobs are constantly being assigned, which causes problems with task cold start. This article suggests a functionality Bayes task recommendation (BTR) method to address the difficulty of the new task suggestion.
    - The crucial concept for handling the dynamism of the crowdsourced system is that BTR understands the implicit factor of the work through task features rather than the task ID and afterwards learns the user’s preference based on their previous actions.
    - Specifically, BTR can not only promptly offer crowdworkers individualised as ask suggestions but also address the task cold-start issue based on the task’s features and the user’s prior behaviour records. Simulations using the actual crowdsourced data set show that BTR outperforms other common strategies that aim to offer the most recent tasks to crowdworkers.

# Background

Recommendation systems are a popular method for providing personalized recom- mendations to users. In the context of research grants, a recommendation system can be used to suggest relevant grant opportunities to researchers, based on their research interests and past performance.

Here are some key concepts and techniques that can be used to develop a recom- mendation system for research grants:

* Collaborative Filtering: Collaborative filtering is a popular technique used in recommendation systems to make personalized recommendations to users based on their past behavior and preferences. In the context of research grants, this could involve analyzing the historical grant data to identify pat- terns and trends in the types of grants that researchers in a particular field tend to apply for.
* Content-based Filtering: Content-based filtering is another technique used in recommendation systems, which involves analyzing the attributes of the items being recommended (in this case, research grants), as well as the char- acteristics of the users. For example, if a researcher has a strong publication record in a particular field, a recommendation system could suggest grants related to that field.
* Hybrid Recommender Systems: Hybrid recommender systems combine both collaborative and content-based filtering techniques to provide more accu- rate and personalized recommendations. By combining these techniques, the recommendation system can take advantage of both the user’s behavior and preferences, as well as the characteristics of the grants being recommended.
* Natural Language Processing (NLP): Natural language processing can be used to analyze the text of grant proposals and identify the research top- ics being proposed. This information can then be used to match researchers with grant opportunities that align with their research interests.
* Machine Learning: Machine learning techniques can be used to train the recommendation system to identify patterns in the historical grant data and to make more accurate predictions about which grants are most likely to be of interest to a particular researcher.
* Evaluation Metrics: It is important to establish appropriate evaluation met- rics to assess the performance of the recommendation system. These may in- clude measures such as precision, recall, and F1 score, which are commonly used in information retrieval and recommendation systems.
* Overall, a recommendation system for research grants would involve ana- lyzing the historical grant data, identifying patterns and trends, and devel- oping appropriate algorithms to make personalized recommendations to re- searchers. The success of the recommendation system would depend on the quality of the data and the algorithms used, as well as the appropriateness of the evaluation metrics used to assess its performance.

# Methodology

## GBDT

* + - A machine learning method called gradient boosting is used, among other things, for classification and regression problems. It provides a predictions model in the form about an ensemble of decisions trees-like weak forecasts. The resultant technique, known as gradient-boosted trees, typically beats random forest when a decision tree is the weak learner. The construction of a gradient- boosted trees model follows the same stage-wise process as previous boosting techniques, but it oversimplifies those techniques by allowing the minimization of any differentiable loss function.
    - Gradient-boosting varies from AdaBoost because it will fit a decision tree also on residuals error (thus the term "gradient") of the previous tree rather than applying weights to particular samples. As a result, rather than explicitly anticipating the target, every successive tree inside the ensembles forecasts the mistake committed by the prior learner.
    - •Decision trees (particularly CARTs) with fixed sizes are utilized generally as base learners when gradient boosting is used. Friedman suggests modifying the gradient boosting methodology for this unique situation in

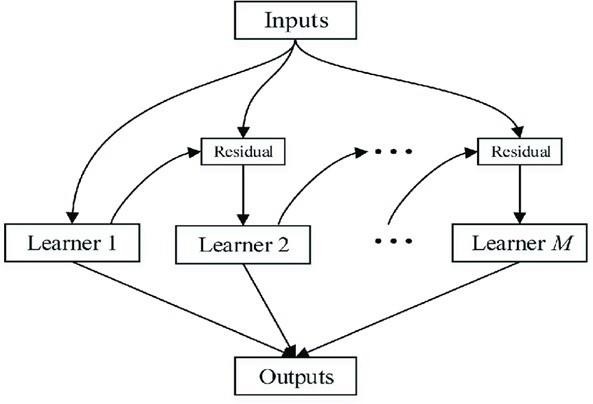
order to improve the level of fit of each base classifier.

* + - Gradient boosting is a technique that may be applied to learning to rank. In their machine-learned ranking algorithms, the commercial search engine Yahoo and Google employ variations of gradient boosting. In High Energy Physics, gradient boosting is also used for data processing. Variants of gradient-boosting Deep Neural Networks (DNN) were effective in recreating the findings of non-machine learning techniques of analysis on datasets used to detect the Higgs boson at the Large Hadron Collider (LHC). In earth and geographical research, gradient-boosted decision trees were also used for example, to assess the quality of sandstone reservoirs.

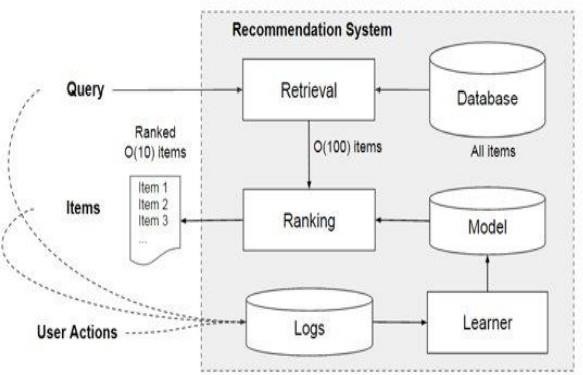
## GBDT application architecture

To use GBDT in the system, you only need to complete the following three steps:

* + - Provide the input data to the GBDT.
    - Learner will give the output and residual will be generated by each learner which will be passed to the other learners.
    - Outputs generated by all the learners will be combined and final output (rec- ommendations) will we generated.



## Recommendation System Architecture



Following are the main building blocks (components) of the recommendation sys- tem:

* + - Database is created manually and with the help of automated web scrapping scripts by searching all over the internet.
    - The system retrieves 100 items from the database when a request is made from the user the retrieved items are them feed to the ranking mechanism.
    - The ranking mechanism uses a two-level layered architecture of its own to rank the items as per the users’ activities and preferences
    - The user is displayed the top 10 items as ranked by the ranking mechanism.
    - The activity of the user is stored as logs and are used by the ML model to further learn about users’ preferences and provide the most personalized ex- perience to the users.

# Results

At the beginning, linear models are a suitable option since they are the simplest to install, interpret, and troubleshoot. Nevertheless, non-linear feature interactions cannot be well modelled by linear models. To harness the potential of our data, we are now using Gradient Boosted Decision Trees (GBDT). The interaction in GBDT models is explicitly represented through a tree structure. GBDT offers a few advantages besides having a bigger hypothesis space, such as managing features with varied ranges, missing feature values, and feature collinearity effectively. Online tests using GBDTs for search ranking were able to produce a statistically significant improvement in the high single-digit percentage range. We are focusing on the following modifications to the GBDT model in order to incorporate search context awareness. In regard to the searcher,

We increased the customizing options. We are expanding the query-candidate matching features for search context. Moreover, and perhaps most significantly, we compared candidates inside the same context, i.e., the same search request, using GBDT models with a pairwise ranking goal. For the same search term, pairwise optimizing compares pairs of impressions. No matter if they are in the identical search query or not, pointwise optimization makes the assumption that each impression is independent. Pairwise ranking is hence more mindful of context. According to online trials, using contextual characteristics and paired GBDT models can boost search engagement metrics by a low two-digit (in the tens) percentage..

# Conclusion

* We are recommending research grants to users according to their interests, and user qualifications, and according to the research grants provider’s needs.
* We compared and study different ML models for the recommendation system we use the GBDT ML model.
* We use web scraping for data generation. We collect the data from different research grants provider’s website and store it in one place and recommend the user research grant.
* At the end, we create UI (User Interface) which will allow even non-technical people to access our work.

# References

* IEEE paper: Personality-Aware Product Recommendation System Based on User Interests Mining and Meta path Discovery
* IEEE paper: Multiple Time Series Perceptive Network for UserTag Suggestion in Online Innovation Community
* IEEE paper: Developing a Meta-Suggestion Engine for Search Queries
* IEEE paper: BTR: A Feature-Based Bayesian Task Recommendation Scheme for Crowdsourcing System
* IEEE paper: BTR: Multiple Time Series Perceptive Network for User Tag Sug- gestion in Online Innovation Community.
* ACM WWW 2019 paper: Entity Personalized Talent Search Models with Tree Interaction Features
* ACM WWW 2019 paper: Entity Personalized Talent Search Models with Tree Interaction Features
* Helpful article: Search Federation Architecture at LinkedIn.
* Helpful article: The AI Behind LinkedIn Recruiter search and recommenda- tion systems