International Journal of Mathematics and Statistics Studies, 13 (1), 1-12, 2025
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Meta-Analysis on the Effectiveness of Personalized Recommendation Systems

Adehi M.U.[1], Erewa-Lawani C.O.[2], Adenomon M.O.[3], Omole O.O.[4], Audi N.I.[5]

[1]Department of Statistics, Faculty of Natural and Applied Sciences, Nasarawa State University, Keffi-Nigeria;
[2]Department of Statistics, Faculty of Natural and Applied Sciences, Nasarawa State University, Keffi-Nigeria;
[3]British Oak Global Institute;
[4]Department of Statistics, Faculty of Natural and Applied Sciences, Nasarawa State University, Keffi-Nigeria;
[5]Department of Statistics, Faculty of Natural and Applied Sciences, Nasarawa State University, Keffi-Nigeria.

doi: https://doi.org/10.37745/ijmss.13/vol13n1112 Published January 02, 2025

Citation: Adehi M.U., Erewa-Lawani C.O., Adenomon M.O., Omole O.O., and Audi N.I. (2025) Meta-Analysis on the Effectiveness of Personalized Recommendation Systems, *International Journal of Mathematics and Statistics Studies,* 13 (1), 1-12

Abstract: *The research study was carried out on meta-analysis on the effectiveness of personalized recommendation systems. Inclusion and exclusion criteria was used to extract the dataset through literature search and article selection. The meta-analysis was based on nine studies consisting of a total of 268,132 observations to conduct a meta-analysis of personalized recommendation systems. The effect size index was the standardized difference mean obtained via a Google search. The random-effects model was employed for the analysis. The studies in the analysis were assumed to be a random sample from a universe of maternal mortality studies in Nigeria. The mean effect size was 1.566 with a 95% confidence interval of 1.194 to 2.053. The Z-value tested the null hypothesis that the mean effect size is 1. We found* $Z = 3.244$ *with* $p < 0.001$ *for* $\alpha = 0.05$*; hence, we rejected the null hypothesis and concluded that the mean effect size was not precisely 1 for personalized recommendation systems. The Q-statistic provided a test of the null hypothesis that nine studies in the analysis share a common effect size; the Q-value is 15.97 with 8 degrees of freedom (k-1) and* $p < 0.001$*. For* $\alpha = 0.1$*, we rejected the null hypothesis that the true effect size was the same in all the 9 studies since Q=k-1, k being the number of studies. The I-squared statistic was 65.3%, which tells us that some 65.3% of the variance in observed effects reflected variance in true effects rather than sampling error. Tau-squared, the variance of true effect sizes, was 0.114 in log units. The study recommended that there should be personalized controlled plans, this will help optimize outcomes and reduce the occurrence of severe mean effects.*
Keyword: meta-analysis, standard difference means, Q-test, personalized recommendation system.

International Journal of Mathematics and Statistics Studies, 13 (1), 1-12, 2025
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

INTRODUCTION

The recent advancements in technology along with the prevalence of online services has offered more abilities for accessing a huge amount of online information in a faster manner. Users can post reviews, comments, and ratings for various types of services and products available online. However, the recent advancements in pervasive computing have resulted in an online data overload problem. This data overload complicates the process of finding relevant and useful content over the internet. The recent establishment of several procedures having lower computational requirements can however guide users to the relevant content in a much easy and fast manner. Because of this, the development of recommender systems has recently gained significant attention. In general, recommender systems act as information filtering tools, offering users suitable personalized content or information. Recommender systems primarily aim to reduce the user's effort and time required for searching relevant information over the internet. Nowadays, recommender systems are being increasingly used for a large number of applications such as web (Castellano et al., 2011), books (Crespo et al., 2011), e-learning (Salehi et al., 2012), tourism (Lorenzi et al., 2011), movies (Bobadilla et al., 2010), music (Yoshii et al., 2008), e-commerce, news, specialized research resources (Porcel et al., 2009), television programs (Shin et al., 2009). While the Internet has gradually become one of the most important and popular sources for people to obtain information and knowledge, it inevitably brings about the problem of" information overload".

LITERATURE REVIEW

Keeping up with the literature of education becomes a more difficult task each year. The Current Index to Journals in Education last year listed more than 17,000 articles published in 700 journals. Research in Education indexed an additional 9000 documents, and Comprehensive Dissertation Abstracts listed more than 6000 dissertations in education. The number of research studies published next year will undoubtedly be greater, and in the year after next, an even larger number of studies is likely to be added to the literature. Researchers have long been aware of the need for organizing this vast literature so that it will be more useful to policy makers, administrators, teachers, and other researchers (Massdex et al, 2001). In recent years some writers have used the term meta-analysis in a broader sense than Glass does. Rosenthal (1984), for example, uses the term to describe almost any attempt to combine or compare statistical results from two or more studies. The other major meta-analytic synthesis of research by Glass and his colleagues was equally impressive (Glass et al., 1982; Smith & Glass, 1980). It focused on the relationship between class size and student learning.

Statistical Developments, the statistical approaches developed in recent times during for combining results from a series of studies were of two types. One approach required researchers to combine probability levels from the studies. The other required researchers first determine whether experiments produced homogeneous results and then make combined estimates of treatment effects. Most methods for combining probability levels are based on a simple fact (Tollerz et al, 1989). If the null hypothesis is true in each study in a set, then p values from statistical tests of all studies will be uniformly distributed between zero and one. That is, the number of outcomes with p values between and effect size.

In 2011, Castellano et al. developed a “NEuro-fuzzy WEb Recommendation (NEWER)” system for exploiting the possibility of combining computational intelligence and user preference for suggesting interesting web pages to the user in a dynamic environment. It considered a set of fuzzy rules to express the correlations between user relevance and categories of pages. Crespo et al., (2011) presented a recommender system for distance education over internet. It aims to recommend e-books to students using data from user interaction. The system was developed using a collaborative approach and focused on solving the data overload problem in big digital content. Lin et al., (2011) have put forward a recommender system for automatic vending machines using Genetic algorithm (GA), k-means, Decision Tree (DT) and Bayesian Network (BN).
(Wang and Wu, 2002) have implemented a ubiquitous learning system for providing personalized learning assistance to the learners by combining the recommendation algorithm with a context-aware technique. It employed the Association Rule Mining (ARM) technique and aimed to increase the effectiveness of the learner’s learning. García-Crespo et al, (2011) presented a “semantic hotel” recommender system by considering the experiences of consumers using a fuzzy logic approach.

METHODOLOGY

In this paper, the method of analysis used in this research work is literature search and articles selection using inclusion and exclusion criteria. A search procedure was executed to find results of empirical studies on personalized recommendation systems. For each study, the following data were extracted: author’s name; publication year; country in which the study was performed; study design; source of recommender system (in content-based recommender systems); online users - sample size; effect size (mean difference); type of outcome (accuracy and effectiveness).

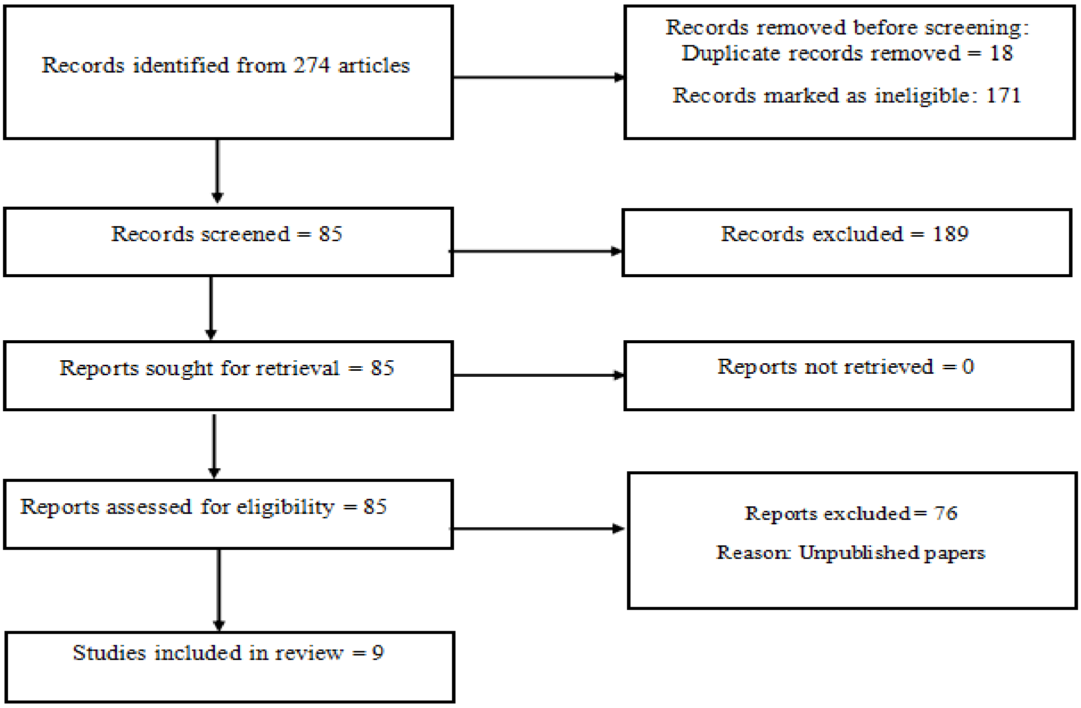


Figure 1: Flowchart diagram showing the inclusion and exclusion process.

The theoretical frame work of this study is based on the assumptions of meta-analysis models. There are fixed and random effect model. (Borenstein et al., 2011)

$$Y_1 = \begin{cases} \vartheta + E_i & fixed\ effect \\ \mu + \vartheta_i + e_i & random\ effect \end{cases}$$

where

E_i and $e_i \sim N(0, \sigma^2)$, where i = 1,2,..,k

Ei is the sampling error
e_i is the random deviations of study's observed effect from the true effect size
ϑ is the population mean
ϑ_i is the true effect size (mean difference)
μ is the grand mean
In a fixed effect analysis, we assume that all the included studies share a common effect size, μ. The observed effects will be distributed about μ, with a variance σ^2 that depends primarily on the sample size for each study.

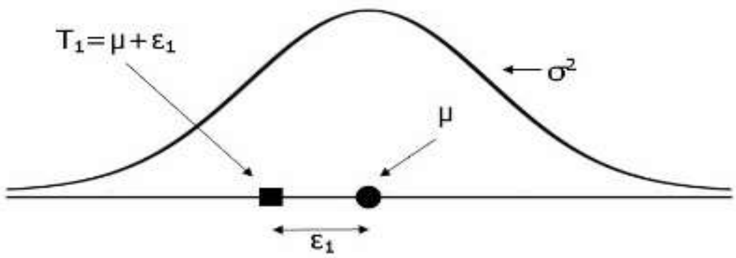


Figure 2: The fixed effect model for the weighted average (Weiss & Daikeler, 2017)

Generally, for any observed effect T_1,

$$T_1 = \mu + \varepsilon_1 \tag{3.0}$$

Assigning weights to the studies

In the fixed effect model, there is only one level of sampling, since all studies are sampled from a population with effect size μ. Therefore, we are dealing with only one source of sampling error-within studies (e).

$$w_i = \frac{1}{\phi_i} \tag{3.1}$$

Where ϕ_i is the within-study variance for study (i).

The weighted mean ($\overline{T}_.$) is then computed as

$$\overline{T_.} = \frac{\sum_{I=1}^{K} W_i T_i}{\sum_{i=1}^{k} w_i} \tag{3.2}$$

That is, the sum of the products $W_i T_i$ (effect size multiply by weight) divided by the sum of the weights.

The variance of the combined effect is defined as the reciprocal of the sum of the weights

$$V_. = \frac{1}{\sum_{i=1}^{k} W_i} \tag{3.3}$$

And the standard error of the combined effect is then the square root of the variance

$$\mathrm{SE}(\overline{T}_.) = \sqrt{V_.} \tag{3.4}$$

The 95% confidence interval will be computed by

$$\text{Lower Limit} = \overline{T}_. - 1.96 * \mathrm{SE}(\overline{T}_.) \tag{3.5}$$

$$\text{Upper Limit} = \overline{T}_. + 1.96 * \mathrm{SE}(\overline{T}_.) \tag{3.6}$$

For the Z value

$$Z = \frac{\overline{T}_.}{\mathrm{SE}(\overline{T}_.)} \tag{3.7}$$

For a one tailed test p-value is

$$P = 1 - \phi(Z) \tag{3.8}$$

and for a two tailed test by

$$p = 2[1 - \phi|z|] \tag{3.9}$$

Where $\phi(z)$ is the standard normal cumulative distribution function.

Random Effect Model

Random effect also called variance component model, is a statistical model where the model parameters are random variables. It is a kind of hierarchical linear model, which assumes that the data being analyzed are drawn from a hierarchy of different populations whose differences relate to that hierarchy.

and computed the variance of these effects sizes (across an infinite number of studies), this variance would be τ^2.

For a set of S effect size measure (γ)

$$\hat{\gamma}_R = \frac{\sum_{i=1}^{S} w_i \hat{\gamma}_i}{\sum_{i=i}^{S} w_i} \tag{3.10}$$

$$W^* = \frac{1}{S^2(\hat{\gamma}_i) + \tau^2} \tag{3.11}$$

$$\tau^2 = \frac{Q - (S-1)}{\sum_{i=i}^{S} w_i - \frac{\sum_{i=i}^{S} w^2{}_i}{\sum_{i=i}^{S} w_i}} \quad \text{for } Q > S-1 \tag{3.12}$$

International Journal of Mathematics and Statistics Studies, 13 (1), 1-12, 2025
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

One method for estimating τ^2 is the method of moments (or the DerSimonian and Laird) method, as follows.

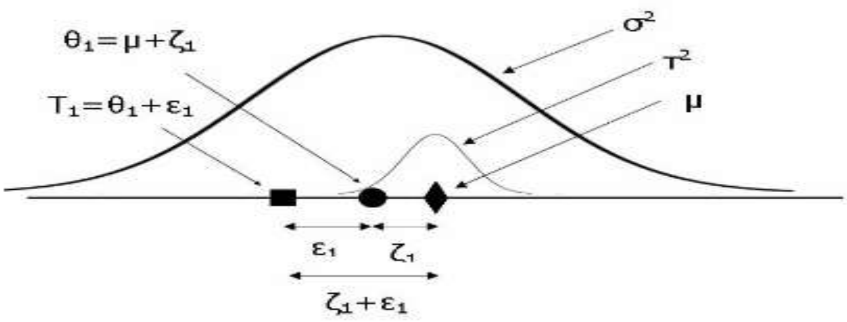


Figure 3: The random effect model for the weighted average (Viechtbauer et al., 2007)

Generally, for any observed effect T_i,

$$T_i = \theta_i + e_i = \mu + \zeta_1 + e_i \qquad (3.13)$$

Assigning weights under the random effects model

In the fixed effect analysis, each study was weighted by the inverse of its variance. In the random effects analysis to each study will be weighted by the inverse of its variance. The difference is that the variance now includes the original (within-studies) variance plus the between-studies variance, tau-squared.

Note the asterisk sign (*) will be used to represent random effect

$$w_i^* = \frac{1}{\phi_i^*} \qquad (3.14)$$

Where ϕ_i^* is the within-study variance for study (i) plus the between-studies variance, tau- squared. That is

$$v_i^* = v_i + \tau^2 \qquad (3.15)$$

The weighted mean ($\overline{T}_.^*$) is then computed as

$$\overline{T}_.^* = \frac{\sum_{i=1}^{k} w_i^* T_i}{\sum_{i=1}^{k} w_i^*} \qquad (3.16)$$

That is, the sum of the products divided by the sum of the weights.

The variance of the combined effect is defined as the reciprocal of the sum of the weights, or

$$v_.^* = \frac{1}{\sum_{i=1}^{k} w_i^*} \qquad (3.17)$$

And the standard error of the combined effect is then the square root of the variance,

$$\mathrm{SE}(\overline{T}_.^*) = \sqrt{V_.^*} \qquad (3.18)$$

The 95% confidence interval for the combined effect is computed as

$$Lower\ Limit^* = \overline{T}_.^* - 1.96^* \ \mathrm{SE}(\overline{T}_.^*) \qquad (3.19)$$

$$Upper\ Limit^* = \overline{T}_.^* + 1.96^* \ \mathrm{SE}(\overline{T}_.^*) \qquad (3.20)$$

For Z- value, could be computed using

$$Z = \frac{\overline{T}_.^*}{\mathrm{SE}(\overline{T}_.^*)}$$

(3.21)

International Journal of Mathematics and Statistics Studies, 13 (1), 1-12, 2025
Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

The one-tailed p-value is area under the probability distribution function (pdf) both to the left of -|z|, and to the right of |z| given by
$p^* = \phi(z)$ and $p^* = 1 - \phi(z)$
(3.22)
And from the fact that $\phi(-z) = 1 - \phi(z)$ the two-tailed p-value by
$p^* = 2[1 - \phi|z|]$
(3.23)
Where $\phi(Z)$ is the standard normal cumulative distribution function

RESULTS AND DISCUSSION

Results of Meta-analysis results to evaluate the effectiveness performance of personalized recommendation systems that is used in online platforms.

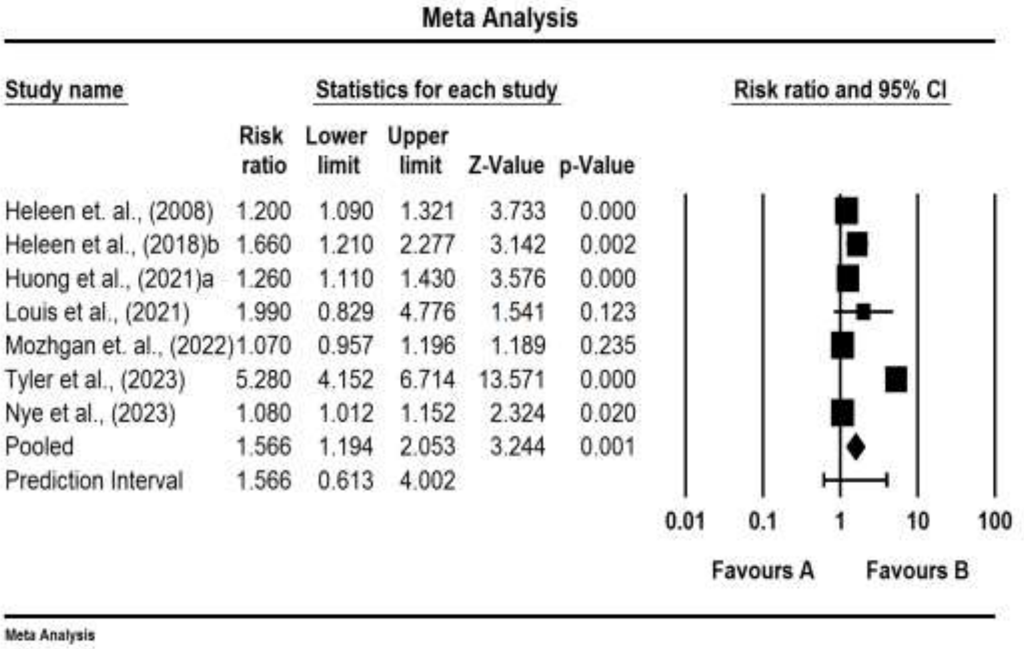


Figure 4: Result of Meta-analysis showing the pooled random-fixed effect model on meta-analysis on the effectiveness performance of personalized recommendation

Print ISSN: 2053-2229 (Print),
Online ISSN: 2053-2210 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

systems that is used in online platform

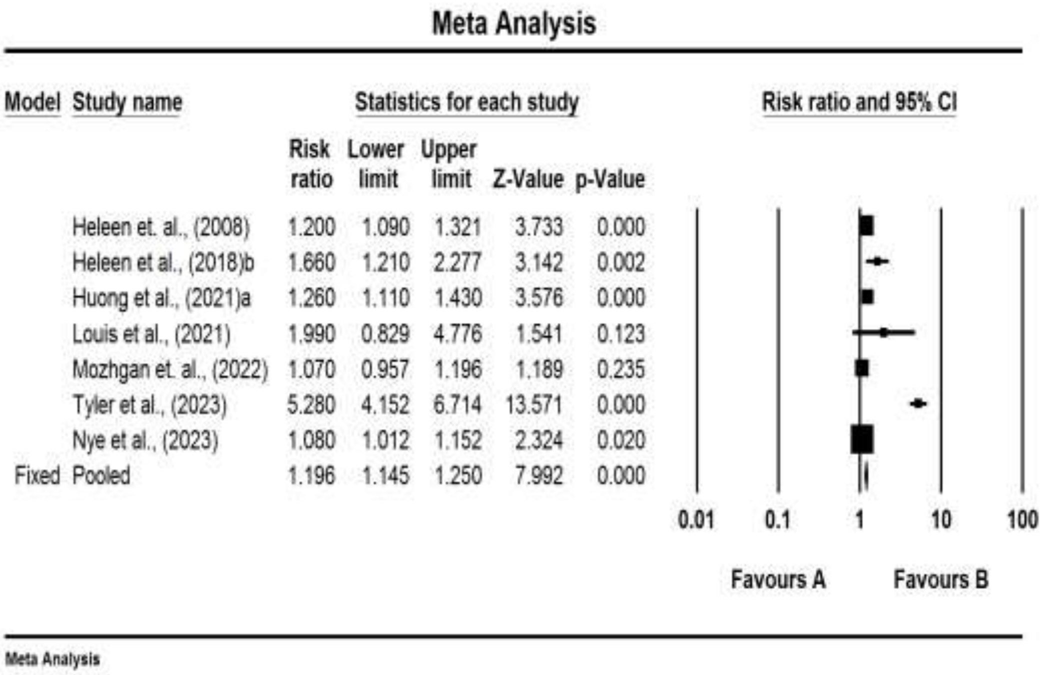


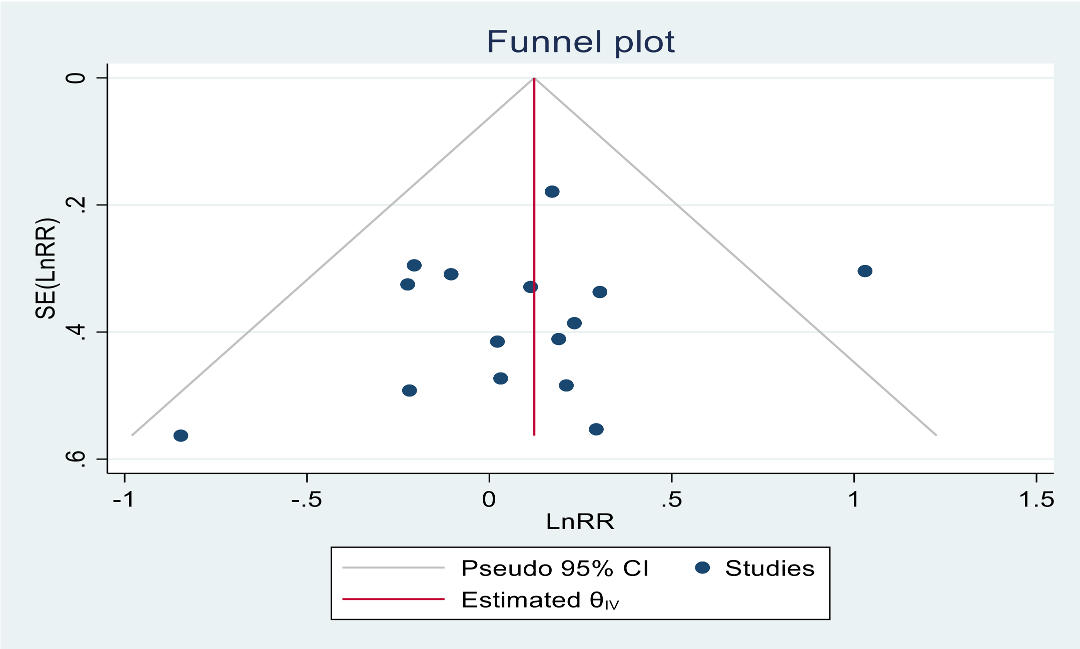
Figure 5: Result of Meta-analysis showing the pooled fixed effect model on meta-analysis on the effectiveness performance of personalized recommendation systems that is used in online platforms.

Interpretation of Meta-analysis result to evaluate the effectiveness performance of personalized recommendation systems that is used in online platforms.

The mean effect size is 1.566 with a 95% confidence interval of 1.194 to 2.053. The mean effect size in the universe of comparable studies could fall anywhere in this interval. The Z-value tests the null hypothesis that the mean effect size is 1.000. The Z-value is 3.244 with p = 0.001. Using a criterion alpha of 0.050, we reject the null hypothesis and conclude that in the universe of populations comparable to those in the analysis, the mean effect size is not precisely 1.000.

The Q-statistic provides a test of the null hypothesis that all studies in the analysis share a common effect size. If all studies shared the same true effect size, the expected value of Q would be equal to the degrees of freedom (the number of studies minus 1). The Q-value is 15.97 with 8 degrees of freedom and p < 0.001. Using a criterion alpha of 0.100, we can reject the null hypothesis that the true effect size is the same in all these studies. The I-squared statistic is 65.3%, which tells us that some 65.3% of the variance in observed effects reflects variance in true effects rather than sampling error. If we assume that the true effects are normally distributed (in log units), we can estimate that the prediction interval is 0.613 to 4.002. The true effect size in 95% of all comparable populations falls in this interval. Tau-squared, the variance of true effect sizes, is 0.114 in log units. Tau, the standard deviation of true effect sizes, is 0.338 in log unit.

Assessing publication bias and testing their symmetry using funnel plot.



Funnel plot of natural unit of standard difference mean and standard error for user behavior for personalized recommendation systems

From 6, the scatter plot of the natural logarithm of effect-size (LnSDM) against their natural logarithm standard errors SE(LnSDM). The estimated effect-size line (LnSDM) and the corresponding pseudo 95% confidence intervals are also plotted. The funnel plot is clearly symmetric, the plotted pseudo confidence interval lines are not genuine confidence interval limits, but they provide some insight into the spread of observed effect-sizes about the estimate of the overall effect-size. From figure 4.2.3 there is no heterogeneity since the studies are scattered within the confidence interval region which resembles an inverted funnel shape, hence there is no publication bias.

CONCLUSION

Given the popularity of streaming services today, research in the field of movie recommendation systems is quite prominent. With new advancements in the field of Artificial Intelligence and Machine Learning, the scope of research on this topic is ever increasing. This research study was aimed at effectiveness that are still prominent in the streaming service recommendation systems and provided logical solutions to those problems by exploring Machine Learning concepts. The primary objective of this thesis was to explore the effectiveness performance of personalized recommendation systems that is used in online platforms and provide the end-user with more personalized results. In this study we have introduced meta-recommenders as a new way to help users find recommendations that are understandable, usable, and helpful. A series of controlled use experiments in the domain of movies indicates that users prefer that these systems provide recommendation data alongside the recommendations and prefer to have control to the selection of this data. Additionally, results suggest that users prefer the recommendations provided by these systems when compared with recommendations provided by "traditional" recommender systems. All told, we feel these results provide a meaningful foundation for the design of future metarecommenders.

Recommendations

Personalized Controlled Plans: Given the variation in controlled response and adverse reactions, personalized controlled plans based on individual patient profiles should be prioritized. This will help optimize outcomes and reduce the occurrence of severe mean effects.

Enhanced Monitoring and Management of Adverse Reactions: Online user providers should implement more rigorous monitoring protocols to manage adverse reactions effectively. This could include pre-controlled evaluations, regular assessments during awareness and follow-ups to mitigate the impact of mean effects.

Policy Implications for Resource Allocation: Policymakers should consider the findings of this study when allocating resources for personalized systems controlled.

Future Studies

This paper is interested in several areas of future work concerning meta-recommenders. These include the transfer of meta-recommenders to other domains, the role of personalization, and real-world acceptance of meta-recommendation systems. While users may not mind providing configuration information to a meta-recommender when the length of the task is relatively short or when encountering a new situation, it is very likely that users will not want to take the time to configure the system for longer or more frequent tasks.

REFERENCES

Adehi M.U and David I.J. Corrections as effect sizes in Meta-Analysis. Science and Technology 2024, 14(1):1-3. DOI: 10.5923/J.scit.20241401.01

Aher SB, Lobo LMRJ. Combination of machine learning algorithms for recommendation of courses in e-learning System based on historical data. Knowl-Based Syst. 2013;51:1–14.

Bakshi S, Jagadev AK, Dehuri S, Wang GN. Enhancing scalability and accuracy of recommendation systems using unsupervised learning and particle swarm optimization. Appl Soft Comput. 2014;15:21–9. 24. Kim Y, Shim K. TWILITE: A recommendation system for twitter using a probabilistic model based on latent Dirichlet allocation. Inf Syst. 2014;42:59–77.

Borenstein, M. (2019). Common Mistakes in Meta-Analysis and How to Avoid Them. Biostat, Inc.

Borenstein, M. (2020). Research Note: In a meta-analysis, the I2 index does not tell us how much the effect size varies across studies. J Physiother, 66(2), 135-139. https://doi.org/10.1016/j.jphys.2020.02.011

DerSimonian, R., & Laird, N. (2015). Meta-analysis in clinical trials revisited. Contemp Clin Trials, 45(Pt A), 139-145. https://doi.org/10.1016/j.cct.2015.09.002

Dong H, Hussain FK, Chang E. A service concept recommendation system for enhancing the dependability of semantic service matchmakers in the service ecosystem environment. J Netw Comput Appl. 2011;34:619–31.

García-Crespo Á, López-Cuadrado JL, Colomo-Palacios R, González-Carrasco I, Ruiz-Mezcua B. Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain. Expert Syst Appl. 2011;38:13310–9.

Gemmell J, Schimoler T, Mobasher B, Burke R. Resource recommendation in social annotation systems: A linearweighted hybrid approach. J Comput Syst Sci. 2012;78:1160–74.

Gottschlich J, Hinz O. A decision support system for stock investment recommendations using collective wisdom. Decis Support Syst. 2014;59:52–62.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Academic Press. Publisher description http://www.loc.gov/catdir/description/els032/84012469.html

Hedges, L. V., & Vevea, J. L. (1998). Fixed and random-effects models in meta-analysis. Psychological Methods, 3(4), 486-504.

Liu L, Xu J, Liao SS, Chen H. A real-time personalized route recommendation system for self-drive tourists based on vehicle to vehicle communication. Expert Syst Appl. 2014;41:3409–17.

Lorenzi F, Bazzan ALC, Abel M, Ricci F. Improving recommendations through an assumption-based multiagent approach: An application in the tourism domain. Expert Syst Appl. 2011;38:14703–14.

Lucas JP, Luz N, Moreno MN, Anacleto R, Figueiredo AA, Martins C. A hybrid recommendation approach for a tourism system. Expert Syst Appl. 2013;40:3532–50.

Mohanraj V, Chandrasekaran M, Senthilkumar J, Arumugam S, Suresh Y. Ontology driven bee's foraging approach based self-adaptive online recommendation system. J Syst Softw. 2012;85:2439–50. 12. Hsu CC, Chen HC, Huang KK, Huang YM. A personalized auxiliary material recommendation system based on learning style on facebook applying an artifcial bee colony algorithm. Comput Math Appl. 2012;64:1506–13.

Niu J, Zhu L, Zhao X, Li H. Afvir: An afect-based Internet video recommendation system. Neurocomputing. 2013;120:422–33.

Rawat YS, Kankanhalli MS. ClickSmart: A context-aware viewpoint recommendation system for mobile photography. IEEE Trans Circuits Syst Video Technol. 2017;27:149–58.

Rice, K., Higgins, J. P. T., & Lumley, T. (2017). A re-evaluation of fixed effect(s) meta-analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society), n/a-n/a. https://doi.org/10.1111/rssa.12275.

Sankar CP, Vidyaraj R, Kumar KS. Trust based stock recommendation system – a social network analysis approach. Procedia Computer Sci. 2015;46:299–305.

Yang S, Korayem M, Aljadda K, Grainger T, Natarajan S. Combining content-based and collaborative fltering for job recommendation system: A cost-sensitive Statistical Relational Learning approach. Knowl-Based Syst. 2017;136:37–45.

Yeh DY, Cheng CH. Recommendation system for popular tourist attractions in Taiwan using delphi panel and repertory grid techniques. Tour Manage. 2015;46:164–76.

Zahálka J, Rudinac S, Worring M. Interactive multimodal learning for venue recommendation. IEEE Trans Multimedia. 2015;17:2235–44.