www.ijramr.com

ISSN : 2350-0743

International Journal of Recent Advances in Multidisciplinary Research

Vol. 08, Issue 03, pp. 6653-6660, March, 2021

RESEARCH ARTICLE

MLM BASED LEARNING & BOOSTING MODEL USING MATCHING USER EVOLUTIVE INTERESTS, COLLABORATIVE DIGITAL RESOURCES AND MICROMETADATA

Ronald Brisebois, *Apollinaire Nadembega and Julien Rault

In Media Technologies, Montréal, Canada

ARTICLE INFO

Article History:

Received 14th December, 2020
Received in revised form
20th January, 2021
Accepted 15th February, 2021
Published online 30th March, 2021

Keywords:

Machine Learning Models, MicroMetadata, Scorm, Semantic Shared Knowledge Notice, User-Created Content, User Interests.

ABSTRACT

Data integration aggregation have allowed to provide uniform interface for multiple heterogonous sources, metadata and MicroMetadata (MM); this issue has attracted a large amount of attention from different areas. Hence, the problemof finding which digital resources may belong to a specific interest demands specific research. We proposed a model named LBAM: The Learning & Boosting Architecture Model. This process makes emphasis on matching user evolutive interests and MM. It combines of context, geolocation, utility, group, content-venue, and user persona aware-approaches. It is a hybrid Machine Learning Model (MLM) and Boosting Models (MLBM): content-based MLM for events semantic MM extraction and collaborative filtering MLM. It uses Machine Learning Models to improve the identification of the User Interests according to different media types. Using simulation study and prototypes, we show that LBAM may propose many personal channels representing slightly the User Interests in a context of aware- approaches. We put in place a first prototype. This paper is the third part of LB project using LBAM.

INTRODUCTION

Crowdsourcing is ingrained in research on open innovation and co-creation and is concerned with whether a wide number of individuals, called "the crowd", can take part actively in a company's innovation processes thereby allowing the company access to intelligence and knowledge that is otherwise dispersed among many users. According to literature, low annotations quality to stem from three possible reasons: (1) unethical spammers submit imprecise or even arbitrary annotations in order to maximize their financial efficiency or due to external distractions; (2) unqualified workers are, despite their best efforts, unable to produce an acceptable annotation quality; (3) malicious workers purposefully aim to undermine or influence the labelling effort; and (4) cognitive biases that are systematic patterns of deviation from norm or rationality in judgment, whereby inferences about other people and situations may be drawn in an illogical fashion [1]. In fact, according to [2], people derive more happiness from anticipating a travel experience than from anticipating possession of something they're going to buy or acquire. In this paper, we proposed a model named LBAM: The Learning & Boosting Architecture Model. This process makes emphasis on matching user evolutive interests and MM. It combines of context, geolocation, utility, group, content-venue, and user persona aware-approaches. It is a hybrid Machine Learning Model (MLM) and Boosting Models (MLBM): content-based MLM for events semantic MM extraction and collaborative filtering MLM. It uses Machine Learning Models to improve the identification of the User Interests according to different media types; this contribution is the following of our previous work [3, 4]. The remainder of the paper is organized as follows. Section 2 presents the related work. Section 3 describes the part 3 of MLM based Learning & Boosting Model (LB) and introduces its various algorithms while Section 4 presents the evaluation through a prototype and a number of simulations. Section 5 presents a summary and some suggestions for future work. The other processes and LBAM architecture will be treated in following papers.

Corresponding author:** ***Apollinaire Nadembega
In Media Technologies, Montréal, Canada.

Related work: Crowdsourcing [1, 5-21] is the practice of obtaining information or input into a task or project by enlisting the services of a large number of people, either paid or unpaid. In [8], crowdsourcing is defined as a branch of co-creation practice that has been made possible through the upsurge of the web, where the 'crowd' can help in validating, modifying and improving a company's value-creating idea or the material it posts over the Internet. According to the literature review, crowdsourcing is an effective way to address such tasks by utilizing a multitude of workers (i.e., the crowd). Crowdsourcing allows solving computer-hard tasks and is benefit data management, such as data cleaning, data aggregation and knowledge identification. Thus, crowd sourced data management has become an area of increasing interest in research and experimentation. Unfortunately, quality control remains one of the main problems.

Eickhoff [1] investigated the prevalence and effect size of a range of common cognitive biases on a standard relevance judgment task; indeed, author strived to demonstrate that cognitive bias can indeed affect crowd sourced labor and leads to significantly reduced result quality. According to him, the common strategies of controlling the crowd by means of qualification tests, demographic filters, incentives, gold standards and sophisticated worker models may not be enough to overcome this new source of noise which is inherently caused by the HIT setup. Author advocated careful task and study design that takes into account cognitive biases to reduce the interface's susceptibility to this kind of label noise. Author observed a balanced tendency of workers following perceived herd behavior. *Author concluded that the more subtle forms of bias, e.g., the Bandwagon or Decoy effects can occur unintentionally in crowdsourcing experiment protocols and should be carefully checked for in order to avoid label degradation.* B. Morschheuser *et al.* [16] investigated how crowdsources' perceived enjoyment and usefulness, behaviors (system usage, crowd sourcing participation, engagement with the gamification feature) and willingness to recommend crowd sourcing approaches are influenced by the use ofcooperative, competitive, and inter-team competitive gamification in crowd sourcing systems. Authors intention is to provide a high external validity; thus, they performed the experiment in the field with a crowdsourcing app called ParKing.

ParKing is a gamified crowd creating system designed to create an interactive map of on-street parking spaces, including the location of parking spaces and their conditions. The gamification component of ParKing attempts to motivate people to participate. ParKing's core game mechanism is the conquering of virtual territories (hexagons) on a map and the constructing of buildings in these territories, visible to the other users of the app. The gameplay is simple; users can earn virtual coins by sharing parking information. These coins can be spent to purchase street segments or construct buildings. Buildings can only be constructed on virtual hexagons, which have been generated and mapped on the real map. *Authors concluded that pure cooperative gamification may not be sufficient to invoke social commitments; in the absence of some external competition or rewards, people in a cooperative setting will perform similarly than when working individually. Thus, authors theorized that inter-team competitive gamification, where users share the goal to win against other teams, may be most effective in invoking cooperative commitments and obligations between users.* A. Ghezzi *et al.* [8] presented an overview of research on crowdsourcing. Authors adopted the Input–Process–Output (I–P–O) framework as the perspective from which to discuss the extant literature, since the process perspective is helpful for integrating different contributions spanning over several theoretical fields. As a result of their study, they are offering an angle for interpreting the extant knowledge and directing future research, achieved by developing a set of suggested research questions. Tong *et al.* [22] proposed a novel data cleaning platform for cleaning multi-version data on the Web, called CrowdCleaner. Crowd Cleaner utilizes crowd sourcing-based approaches for detecting and repairing errors that usually cannot be solved by traditional data aggregation. According to authors, Crowd Cleaner does not only detect and repair false or delay versions of updates but also automatically determines which version of data should be accepted. *Unfortunately, authors do not demonstrate clearly the performance of their Crowd Cleaner.* M. Li *et al.* [15] conceptualized a blockchain-based decentralized framework for crowdsourcing named Crowd BC, in which a requester's task can be solved by a crowd of workers without relying on any third trusted institution, users' privacy can be guaranteed and only low transaction fees are required. Crowd BC aims is to design a decentralized crowdsourcing system with reliability, fairness, security and low services fee. They introduced the architecture of their proposed framework, based on which we give a concrete scheme. Crowd BC uses a smart contract to perform the whole process of crowdsourcing task which contains task posting, task receiving, reward assignment, etc. The smart contracts include: User Register Contract (URC), User Summary Contract (USC), Requester-Worker Relationship Contract (RWRC), by which crowdsourcing functionalities can be achieved such as posting and receiving a task without relying on any central authority. Authors implemented the proposed scheme to verify the feasibility through a software prototype based on Ethereum public test network. *Unfortunately, authors do not present specifically the cold-start strategy for the User Register Contract of smart contracts component.* S. Paun *et al.* [21] proposed a mention pair-based approach to aggregating crowd sourced anaphoric annotations.. They introduced a mention pair-based approach to aggregating crowd sourced anaphoric annotations and assessed the quality of the inferred pairs, of the post-hoc constructed co-reference chains, and the viability of using the inferred chains as an alternative to gold chains when training a state of the art co-reference system. *In the mention pair model, the task of linking the mention to a co-reference chain/entity is split in two parts: classifying mention pairs as co-referring or not, and subsequent clustering. Unfortunately, authors contribution just addresses the first part of the model. As conclusion, we can claim the most of existing approaches are based on the crowd sourced based on the workers who answer about the similarity between pair entities. We also understand that the best approach is one that uses at the least human contribution while achieving high accuracy. In addition, gamification and reward are good way to motivate users to create or generate content. Finally, no approach shows how their ERM is used to update the entity repository.*

Personal Agenda and Channels Portal (PACP) using Matching User Evolutive Interests (MUEI), Personal Gaming and Learning (PGL), Personal Digital Resources (PDR) and Personal Secured MicroMetadata Space (PSMS). In this section, we present the details of the proposed approach using SMESE. First, we introduce MLM based Learning & Boosting Model and second, some details of LBAM algorithms and models (Part 3). For further understanding about SMESE algorithms and processes to semantically enrich metadata using multiple metadata/data sources, refer to previous papers [23]. The Life Booster project proposed to use the SMESE platform to create User Evolutive Interests, 3 portals (Personal Agenda & Channels, Collaborative Learning & Events, Collaborative Digital Resources) and 1 Personal User Space – see Fig. 1.

Fig. 1. LB project outputs

According to [24], existing literatures on ERS ignore the social aspect of events; indeed, people prefer to attend events with their friends or family rather than alone. In this research work, the proposed ERS model, call Users/Events Social Matching Model, is being designed to take into account the social aspect of events in order to provide happiness to individual user and groups of users such as family or colleagues. MEUI aims is to match users interests and dynamic personas (Personal DNA) with events semantic metadata and hidden characteristics taking into account: (1) context, (2) geolocation, (3) utility, (4) group, (5) content-venue and (6) emotion and sentiment aware approaches. MEUI is a Hybrid Machine Learning Model (HMLM) that used content-based MLM for events semantic and hidden metadata extraction and collaborative filtering MLM for user personas learning. The MEUI architecture is divided into three modules: (i) data collection which extracts the unstructured dataset from the several event-based social networks (EBSNs) such as Facebook using API's; (ii) data mapping module which is basically used to integrate the common knowledge/data that can be shared between considered different EBSNs. This module integrates and reduces the data into structured events' instances. As the dataset was collected from more than ten different sites, an intersection of all was taken out. This two modules are proposed in our previous work [25, 26] which are based on [27-35]; and (iii) MEUI algorithm for users and events matching who is part of our project [36].

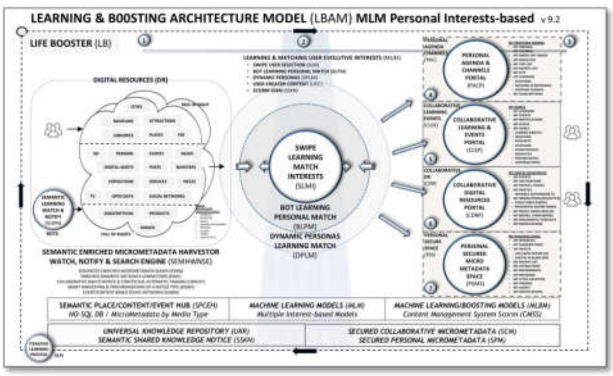

Fig. 2: LBAM Overview Model

Overview of Life Booster project: The Fig. 2 represents the Learning & Boosting Architecture Model (LBAM), a Machine Learning Interest-based Model, has goals: 1) to identify Matching Evolving. User Interest (MEUI) of person and 2) potentially to boost daily their life by providing to them a proposed Personal Agenda and Channels according to a set of Personal Metrics (PM) and interests who evolve periodically. This LBAM model is built from 3 main processes: a) Identification of the MicroMetadata (MM) as MicroContent of Digital Resources (DR) including Events and their timeline (novelties) and ongoing enrichment; b) MEUI using a Bot and a swipe action; and c) The Daily Smart Booster Agenda created to suggest DR according to the evolutive user interests. This project Booster (LB) intends to keep track of the rights of the contents (Digital Resources) or Events, the MEUI and MLBM who are part of an Iterative Learning Process (IPL) shown in the Fig. 3.

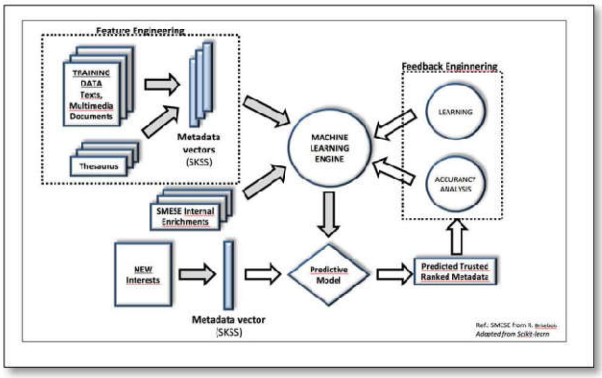

Fig. 3: Iterative Learning Process (IPL)

The **first process** is based on the creation of a hub of secured multiple metadata using the Semantic Enriched MM Harvestor, Watch, Notify & Search Engine linked to Users and Bots (SLWN) and including multiple sources of rights and their aggregation into MM by Media Type (Notice Type as City, Museum, Place, Event, Person, etc.) using Multi Sources Semantic Knowledge (SSKN). These SSKN could be enriched enriched to create Enriched MM. These Metadata are assembled through a Harvesting process able to catalogue the Rights, the Interests and the Novelties. This process includes Sub-processes named: Federated Enriched MM Search (FEMS), Enriched Semantic Metadata Connectors (ESMC), Collaborative Rights Notice & Contextual Automatic Tagging (CRNCAT), Smart Harvesting & Synchronization of a Notice

Type (SHSNT) and Event/Content-based Social Network (ECBSN). This process includes too the ability for the User or the Merchant to create or update media and metadata. This harvesting process has to keep track too of the Novelties. SLWN allows to keep track of any event who may interest some watching and notifying process in the system. SEMHWNS includes the following main process: Federated Enriched MM Search (FEMS), Enriched Semantic Metadata Connectors (ESMC) – (Enrichments are per examples: Interests, Novelties, Persons, etc.), Collaborative Rights Notice & Contextual Automatic Tagging (CRNCAT), Smart Harvesting & Synchronization of a Notice Type and Event/Content-Based Social Networks (ECBSN). This process harvests Free of right and Full of Right Content and manage the MM multi-rights.

The **second process** is mainly to identify the MEUI by an Algorithm of matching from four different levels of User Interests: a) The User Personal Interest using the real time Swipe Learning Match Interests (SLMI); b) The Interests of the Personas of the User using Dynamic Personas Learning Match (DPLM) – the Personas of the Users are categorized in 18 different personas in our model; c)

The Bot swipe as a counterpart for Swipe Learning Match Interests (SLMI) using Bot Learning Match (BLM) – a simulator of automatic matching interests based on a set of user with the 95% of the same Personas and; d) User Created Content (UCC) allowing to extract some behavior from the User. The Bot Learning Match (BLM) is an assisted process (ChatBot) allows to match User Interests for Digital Assets as Events, Photos, Persons, etc. This process uses Multiple Interest-based Models to learn the User Interests in different situations with the Swipe principle to like (right) or to don't like (left), time of the day and contextual behavior. Using MLM, this process improves the MEUI identification over the learning process.

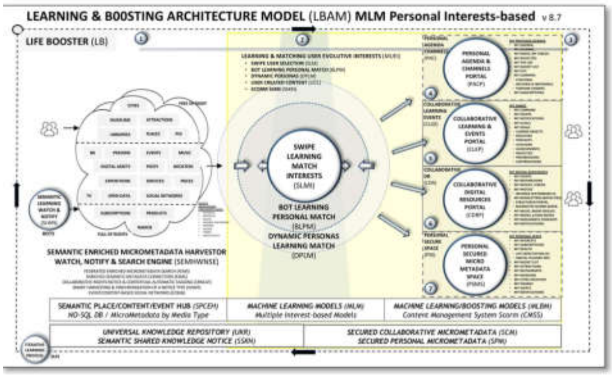

Fig. 4. The process three to seven (in yellow)

The third process (Fig. 4) focus on the prediction of the daily evolving interests of each user and context regarding: Personal Agenda and Channels Portal – it is a personal Journal, a personal Radio and a personal TV channel (PACP). Here we build a recommended agenda, journal, radio channel and videos channel to a specific user according to the entire five process of LBAM and his evolving personal interests. PACP propose to the User an Agenda for the day or the coming week, every day this agenda is refined according to the usage and interests of the users. This process uses Machine Learning/Boosting Models to: a) improve the cataloguing of the Digital Asset and Events; b) to boost interest of User and c) to improve the identification of the User Interests. This process makes emphasis too on Collaborative Learning & Events Portal (CLEP) gives games or learning activities to do. The Collaborative Digital Resources Portal – Collaborative Digital Resources identifies potential Events and Media who could meet the Interests of the User. And the last process is Personal Secured MM Space (PSMS), where the user can manage his Configuration, Interests, Digital Resources, Events and Agenda and regroup all personal information. This Space includes too Personal Metrics (PM) and Security. The fourth process is the PACP Process but with an emphasis on Personal Channels (PC) process. It allows to propose to User a dedicated Personal Channel according to his interests and available Digital Resources at a specific time. This Personal Agenda & Channel Portal is using MLBM evolving with time and all interactions with the User. The fifth process named Collaborative Learning & Events Portal (CLGP) includes the sharing of knowledge and gaming for the benefice of every user. The process includes the ability to create, reference, evaluate and organize content or knowledge in a evolutive learning process at different level. It allows Digital Resources to be accessed and used by a multitude of user in many languages. The sixth process is the Collaborative Digital Resources Portal (CDRP). The process includes My Newsletter who fulfill the CDRH to create content and digital resources per different interest categories and learning needs. This process includes too a CMS based Micro-Sites Generator using newsletter smart aggregation to create new content and knowledge. This process includes notifications and alerts according to the interests of the users, it is called Watch for me. The seventh process is the Secured Personal MM Space (SPMS) but with an emphasis on Personal Metrics (PM) and Digital Placebo (DP). The process includes in My Health, the Life expectation metric and the DP who intend to help User to reach a better level on MEUI. All these seven processes are embedded in a larger Machine Learning Mechanism allowing to learn at different stages of the macro process and to improve all other learning processes. All process used a multilingual thesaurus. We call this critical process: Iterative Learning Process (ILP). The remainder of the paper is organized as follows. Section 2 presents the related work. Section 3 describes the part 3 of MLM based Learning & Boosting Model (LB) and introduces its various algorithms while Section 4 presents the evaluation through a prototype and a number of simulations. Section 5 presents a summary and some suggestions for future work.

Algorithms

The following Fig. 5 presents at a high level the algorithm to map User' Interests with DR, Events and Enriched Micro Metadata (Micro Content). Usage Affinities are identified by many mechanisms at the daily personal usage of the Life Booster application.

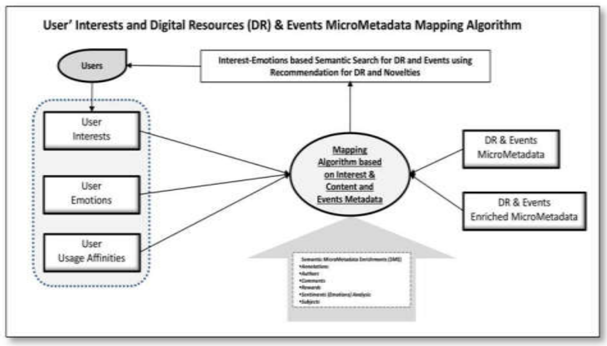

Fig. 5. MicroMetadata (MM) Mapping Algorithm

Machine learning model (MLM) and Museum Prototype: MLM algorithms are used at different levels in LBAM to identify the evolutive interests of users. It uses the same model than SMESE but enhances the process to identify sources, rights and MM's enrichment in the structured environment and unstructured web, see Fig. 6.

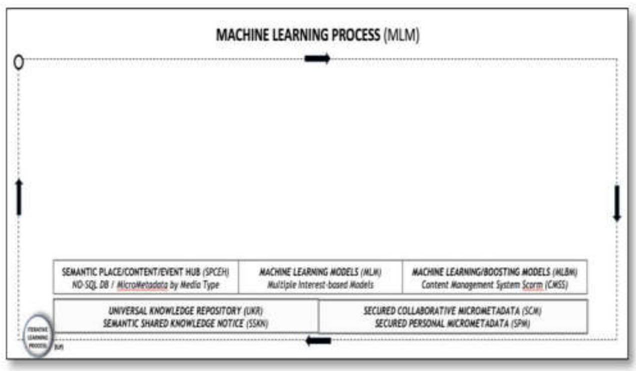

Fig. 6. Machine Learning Models

In Fig. 7 we can see the prototype for Museum, this APPS allows to harvest MM linked to the SSKN of Museum and to harvest MM from the structured and unstructured web.

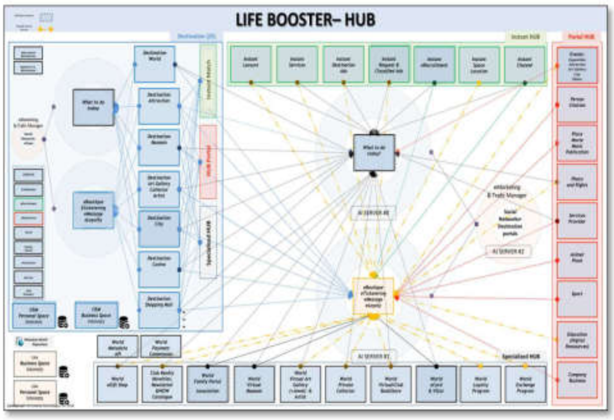

Fig. 7: Life Booster - Hub

Prototype Applications and Performance Evaluation: Prototype: Learning Application (APPS)

The prototype is an APPS and a portal allowing user to select his interests and allowing to test our algorithm to refine and identify user's evolutive interest. To achieve this goal, we need some accuracy and precision evaluation for:

- Entity resolution;
- Content linkage;
- Micro-metadata rights preservation.

In the Fig. 8 and Fig. 9, we present the structure of the APPS and the prototype.

Simulation Setup and Datasets Characteristics

The datasets we use was provided by our Contents Universal Repositories which contains 267 different types of real entities and more than 100 million entities. The contents are been harvested using our previous contributions process [25, 26]. For example, our Contents Repositories consist of: Museum, PointOfInterest, Artwork, Artist, Video, Movie, Music, Radio, TV, Bar, Restaurant, Cinema, Theater, Event, Place, Organization, Person, Restaurant, MovieClip, RadioClip, TVClip, VideoGameClip, PodcastSeries, RadioSeries, TVSeries, VideoGameSeries, BookSeries, MovieSeries, MusicPlaylist, Painting, Photograph, on so on.

Performance measurement criteria. The goal of this section is to compare the performance of three different approaches with our proposal, called LBAM-MLBM. To evaluate the behavior of comparison approaches, we employ two kinds of measures: Number of Recommendations when varying the number of users and the Boosting Level when varying the number of different types au contents suggested to users. The criterion "Number of recommendations" denotes the average number of times recommendations which were made before the satisfaction of a given user. The criterion "Boosting Level" denotes the average number of users happy after accepted a recommendation. As comparison terms, we use the approaches described in [37], [38], and [39], which are referred to as MLM_1, MLM_2, and MLM_3, respectively.

RESULTS AND DISCUSSION

In Fig. 10, we evaluate the average number of recommendations made before user satisfaction; for example, for LBAM-MLBM, only one recommendation is required to satisfy each user in a group of ten users. We observe that, for all the models, the number of recommendations before satisfactions increases with the number of users; this is expected since when the number of users increases, the number of personal interests increases and thus the average number of recommendations before satisfaction increases. We also observe in Fig. 10 that, LBAM-MLBM outperforms MLM_1, MLM_2 and MLM_3; for example, LBAM-MLBM provides an average of 2.8 number of recommendations before the satisfaction of a group of ten users, whereas MLM_3 (more efficient than MLM_1 and MLM_3 in this scenario) provides an average of 5.5 number of recommendations before the satisfaction of a group of ten users; overall, the average relative improvement of LBAM-MLBM compared with MLM_3 is about 2.7 number of recommendations before the satisfaction of a group of ten users; this means that, LBAM-MLBM needs 2.7 number of recommendation less than MLM_3 to satisfya group of ten users. This can be explained by the fact that LBAM-MLBM combines of context-aware approach, geolocation-aware approach, utility-aware approach, group-aware approach, content-venue-aware approach, and user persona-aware approach.

In Fig. 11, we evaluate the average number of users happy after accepting the recommendations performed by the MLM algorithms according to the diversity of the types of content available; for example, in the repository of five different types of content, 71% of users become happy when they accept LBAM-MLBM recommendations. We also observe in Fig. 11 that, LBAM-MLBM outperforms MLM_1, MLM_2 and MLM_3; for example, LBAM-MLBM provides an average of 0.78 of users Boosting Level using five different contents types, whereas MLM_3 (more efficient than MLM_1 and MLM_3 in this scenario) provides 0.65 of users Boosting Level using five different contents types; overall, the average relative improvement of LBAM-MLBM compared with MLM_3 is about 13% of users Boosting Level using five different contents types; this means that, for the same number of different contents types, LBAM-MLBM provides 13% more happiness than MLM_3. This can be explained by the fact that LBAM-MLBM algorithms uses Enriched Semantic Micro-Metadata instead of metadata to identify the MEUI based on four different levels of User Interests: (i) User Personal Interest using the real time Swipe Learning Match (SLM), (ii) Interests of the Personas of the User

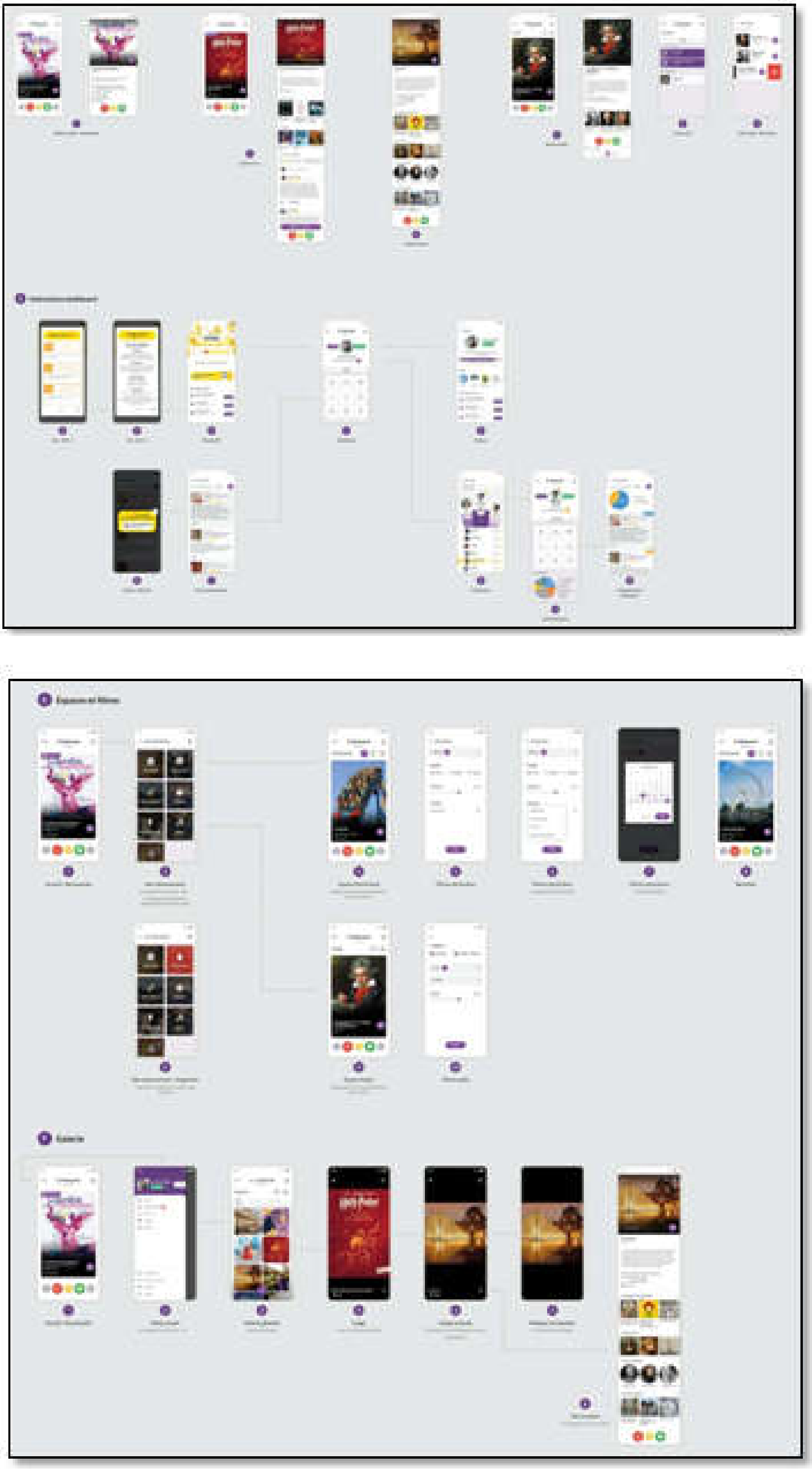

Fig. 9. The organization of the mobile APPS (2/2)

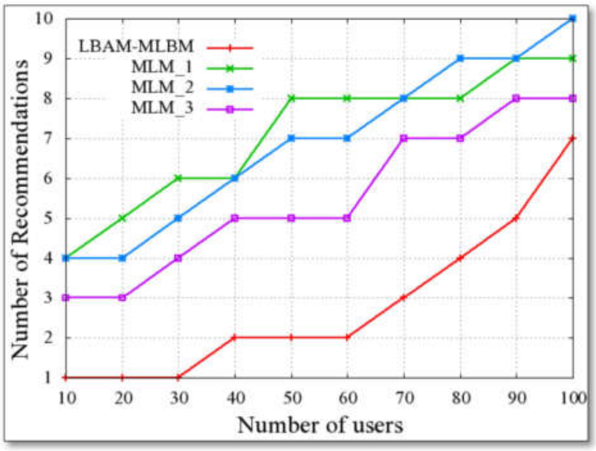

Fig. 10: Number of Recommendations before satisfaction Vs Number of Users

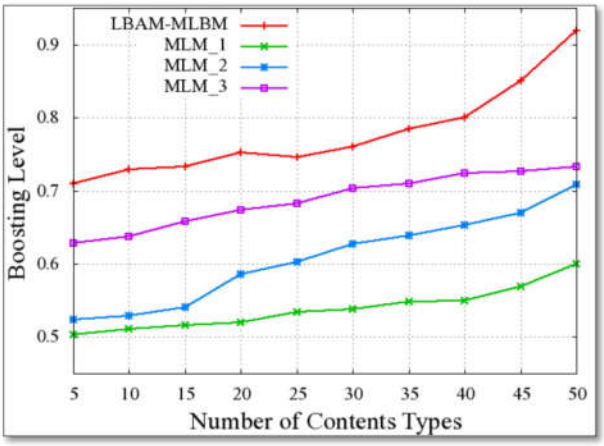

Fig. 11. Boosting Level Vs Number of Content Types

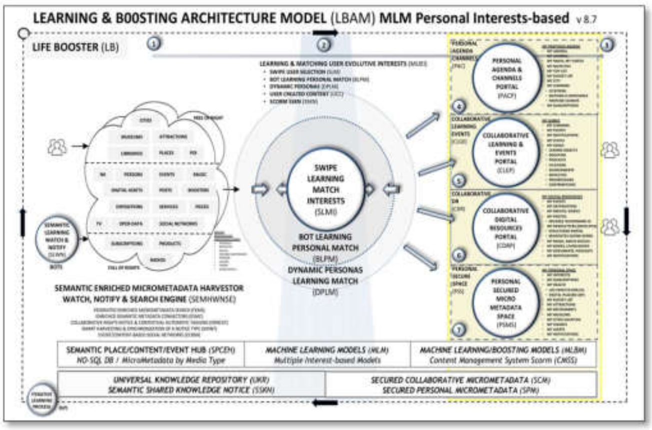

Fig. 12. The four to seven (in yellow)

using Dynamic Personas Learning Match (DPLM) and (iii) Bot swipe as a counterpart for Swipe Learning Match (SLM) using Bot Learning Match (BLM).

Summary and future work

In summary, the analysis of the simulation results shows also that schemes that use human annotation combine to MLM outperform schemes that are limited to human annotation or machine learning model.

We also observe that schemes that use MLM outperform schemes that are limited to human annotation. Yet, there many improvements that can be added to this model: refinements of SSKN, SLM and BLM.Here (see Fig. 12) are some of the future work that we looking to explore furthermore;this is the fourth to seventh process of LBAM.

REFERENCES

[1] Eickhoff, C. "Cognitive Biases in Crowdsourcing," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Marina Del Rey, CA, USA, 2018, pp. 162–170.

[2] A. Kumar, M. A. Killingsworth, and T. Gilovich, "Waiting for Merlot: Anticipatory Consumption of Experiential and Material Purchases," *Psychological Science,* vol. 25, no. 10, pp. 1924–1931, 2014.

[3] R. Brisebois, A. Nadembega, T. Hajj, and C. Charles, "Mlm-based learning and boosting model – part 1: multi-sources/rights of digital resources to build universal knowledge repositories using an enriched semantic micro metadata harvester engine and semantic shared knowledge notice," *Asian Journal of Science and Technology (AJST),* vol. 11, no. 01, pp. 10722-10733, Jan. 2020, 2020.

[4] R. Brisebois, and A. Nadembega, "MLM based Learning & Boosting Model – Part 2: Matching User Evolutive Interests (MUEI) based on Multi-Engines Learning Match Algorithms (ELMA) using User Created Content and Universal Knowledge Repositories (UKR)," *International Journal of Engineering Science Invention (IJESI),* vol. 09, no. 11, pp. 50-64, Apr. 2020, 2020.

[5] S. Chawla, J. D. Hartline, and B. Sivan, "Optimal crowdsourcing contests," *Games and Economic Behavior,* vol. 113, pp. 80-96, 2019/01/01/, 2019.

[6] X. Cheng, Q. Gou, J. Yue, and Y. Zhang, "Equilibrium decisions for an innovation crowdsourcing platform," *Transportation Research Part E: Logistics and Transportation Review,* vol. 125, pp. 241-260, 2019/05/01/, 2019.

[7] F. Daniel, P. Kucherbaev, C. Cappiello, B. Benatallah, and M. Allahbakhsh, "Quality Control in Crowdsourcing: A Survey of Quality Attributes, Assessment Techniques, and Assurance Actions," *ACM Comput. Surv.,* vol. 51, no. 1, pp. 1-40, 2018.

[8] A. Ghezzi, D. Gabelloni, A. Martini, and A. Natalicchio, "Crowdsourcing: A Review and Suggestions for Future Research," *International Journal of Management Reviews,* vol. 20, no. 2, pp. 343-363, 2018.

[9] G. V. Horn, S. Branson, S. Loarie, S. Belongie, and P. Perona, "Lean Multiclass Crowdsourcing." pp. 2714-2723.

[10]E. Segev, "Crowdsourcing contests," *European Journal of Operational Research*, 2019/03/06/, 2019.

[11]K. B. Sheehan, "Crowdsourcing research: Data collection with Amazon's Mechanical Turk," *Communication Monographs,* vol. 85, no. 1, pp. 140-156, 2018/01/02, 2018.

[12]Y. Tong, L. Chen, Z. Zhou, H. V. Jagadish, L. Shou, and W. Lv, "SLADE: A Smart Large-Scale Task Decomposer in Crowdsourcing," *IEEE Transactions on Knowledge and Data Engineering,* vol. 30, no. 8, pp. 1588-1601, 2018.

[13]I. Blohm, S. Zogaj, U. Bretschneider, and J. M. Leimeister, "How to Manage Crowdsourcing Platforms Effectively?," *California Management Review,* vol. 60, no.2, pp. 122-149, 2018.

[14]J. Le Coz, A. Patalano, D. Collins, N. F. Guillén, C. M. García, G. M. Smart, J. Bind,

A. Chiaverini, R. Le Boursicaud, G. Dramais, and I. Braud, "Crowd sourced data for flood hydrology: Feedback from recent citizen science projects in Argentina, France and New Zealand," *Journal of Hydrology,* vol. 541, pp. 766-777, 2016/10/01/, 2016.

[15]R. Martella, E. Clementini, and C. Kray, "Crowdsourcing geographic information with a gamification approach," *Geodetski vestnik,* vol. 63, pp. 213-233, 06/01, 2019.

[16]B. Morschheuser, J. Hamari, and A. Maedche, "Cooperation or competition – When do people contribute more? A field experiment on gamification of crowdsourcing," *International Journal of Human-Computer Studies,* vol. 127, pp. 7-24, 2019/07/01/, 2019.

[17]J. Wang, J. Zhao, Y. Zhang, X. Peng, Y. Li, and Y. Xie, "Enabling Human-Centric Smart Cities: Crowdsourcing-Based Practice in China," *Computer,* vol. 51, no. 12, pp.42-49, 2018.

[18]R.-Q. Wang, H. Mao, Y. Wang, C. Rae, and W. Shaw, "Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data," *Computers & Geosciences,* vol. 111, pp. 139-147, 2018/02/01/, 2018.

[19]J. Bragg, Mausam, and D. S. Weld, "Sprout: Crowd-Powered Task Design for Crowdsourcing," in Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, Berlin, Germany, 2018, pp. 165-176.

[20]F. Palomba, M. Linares-Vásquez, G. Bavota, R. Oliveto, M. D. Penta, D. Poshyvanyk, and A. D. Lucia, "Crowdsourcing user reviews to support the evolution of mobile apps," *Journal of Systems and Software,* vol. 137, pp. 143-162, 2018/03/01/, 2018.

[21]S. Paun, J. Chamberlain, U. Kruschwitz, J. Yu, and M. Poesio, "A Probabilistic Annotation Model for Crowdsourcing Coreference." pp. 1926-1937.

[22]Y. Tong, C. C. Cao, C. J. Zhang, Y. Li, and L. Chen, "CrowdCleaner: Data cleaning for multi-version data on the web via crowdsourcing." pp. 1182-1185.

[23]A. Jurek, J. Hong, Y. Chi, and W. Liu, "A novel ensemble learning approach to unsupervised record linkage," *Information Systems,* vol. 71, pp. 40-54, 2017/11/01/, 2017.

[24]H. Yin, L. Zou, Q. V. H. Nguyen, Z. Huang, and X. Zhou, "Joint Event-Partner Recommendation in Event-Based Social Networks." pp. 929-940.

[25]R. Brisebois, A. Nadembega, and T. Hajj, "Traceable and trusted smart harvesting algorithm from unstructured and structured web (smese-ttsha)," *International Journal of Current Research (IJCR),* vol. 11, no. 02, pp. 1050-1058, February, 2019, 2019.

[26]R. Brisebois, A. Nadembega, and T. Hajj, "Trusted smart harvesting algorithmbased on semantic relationship and social networks (SMESE-TSHA)," *International Journal of Recent Scientific Research (IJRSR),* vol. 10, no. 01, pp. 30593-30604, January,, 2019.

[27]R. Brisebois, A. Abran, A. Nadembega, and P. N'techobo, "An Assisted Literature Review using Machine Learning Models to Recommend a Relevant Reference Papers List," *International Scientific Research Organization Journal,* vol. 02, no. 02, pp. 1- 24, November 2017, 2017.

[28]R. Brisebois, A. Nadembega, P. N'techobo, and H. L. Djeuteu, "A semantic web metadata harvesting and enrichment model for digital library and social networks," *International Journal of Current Research (IJCR),* vol. 9, no. 10, pp. 59162-59171, October 2017, 2017.

[29]R. Brisebois, A. Abran, A. Nadembega, and P. N'techobo, "An Assisted Literature Review using Machine Learning Models to Identify and Build a Literature Corpus," *International Journal of Engineering and Science Invention (IJESI),* vol. 6, no. 7, pp.72-84, 2017.

[30]R. Brisebois, A. Abran, A. Nadembega, and P. N'techobo, "Efficient Scientific Research Literature Ranking Model based on Text and Data Mining Technique," *International Journal of Engineering Research And Management (IJERM),* vol. 04, no. 02, pp. 95-105, February 2017, 2017.

[31]R. Brisebois, A. Abran, A. Nadembega, and P. N'techobo, "A Semantic Metadata Enrichment Software Ecosystem based on Machine Learning to Analyse Topic, Sentiment and Emotions," *International Journal of Recent Scientific Research (IJRSR),* vol. 8, no. 4, pp. 16698-16714, April,, 2017.

[32]R. Brisebois, A. Abran, and A. Nadembega, "A Semantic Metadata Enrichment Software Ecosystem (SMESE) Based on a Multi-Platform Metadata Model for Digital Libraries," *Journal of Software Engineering and Applications (JSEA),* vol. 10, no. 04,pp. 370-405, April 2017, 2017.

[33]R. Brisebois, A. Abran, A. Nadembega, and P. N'techobo, "Text and Data Mining & Machine Learning Models to Build an Assisted Literature Review with Relevant Papers," *International Journal of Scientific Research in Information Systems and Engineering (IJSRISE),* vol. 03, no. 01, pp. 6-27, April 2017, 2017.

[34]R. Brisebois, A. Abran, A. Nadembega, and P. N'techobo, "A semantic metadata enrichment software ecosystembased on machine learning to analyze topic, sentiment and emotions," *International Journal of Recent Scientific Research (IJRSR),* vol. 08, no. 04, pp. 6698-16714, April 2017, 2017.

[35]R. Brisebois, A. Abran, and A. Nadembega, "A Semantic Metadata Enrichment Software Ecosystem (SMESE) based on a Multi-platform Metadata Model for Digital Libraries," *Journal of Software Engineering and Applications (JSEA),* vol. 10, pp. 370-405, April 30, 2017.

[36]R. Brisebois, A. Nadembega, J. Rault, and W. Hoffman, "LUCY: A Smart City Mobile Application Architecture Model for Multidimensional Interests Spaces and Personas using Machine Learning Models," *Asian Journal of Science and Technology (AJST),* vol. 10, no. 09, pp. 10067-10078, September, 2019, 2019.

[37]J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, "Collaborative filtering and deep learningbased recommendation system for cold start items," *Expert Systems with Applications,* vol. 69, pp. 29-39, 2017/03/01/, 2017.

[38]A. S. Tewari, J. P. Singh, and A. G. Barman, "Generating Top-N Items Recommendation Set Using Collaborative, Content Based Filtering and Rating Variance," *Procedia Computer Science,* vol. 132, pp. 1678-1684, 2018/01/01/, 2018.

[39]G. D. S. P. Moreira, D. Jannach, and A. M. D. Cunha, "Contextual Hybrid Session- Based News Recommendation With Recurrent Neural Networks," *IEEE Access,* vol. 7, pp. 169185-169203, 2019.

****