

Rice University's Technology Development Fund.

Proposal submission at Rice portal [DEADLINE: October 30, 2021.**]**

- ☐ A single-page budget indicating how the requested funds will be expended;
- ☐ A curriculum vitae for each applicant of no more than three pages per CV.

Project Description. -Three (3) pages max-

- ☒ the scope of the proposal
- ☒ the relevant market need and size
- ☒ the commercial potential of the proposal
- ☒ a probable development timeline for productization
- ☒ identifying the potential product or products to be created
- ☒ income-generation potential associated with market entry
- ☒ any interest expressed in the project by potential industry partners or commercial sponsors
- ☒ any relevant patent or technology disclosures on file with the Office of Technology Transfer.

1 The societal need and the customer

Small and medium businesses (SMB) worldwide are currently looking to improve their online operations to serve a market shifting to buy mainly on the web, which is often an overlooked collateral effect of the past stay-at-home restrictions worldwide. Recommendation systems are a valuable tool in e-commerce for SMB because they reduce data overload by providing meaningful results to customers that increase the possibility of converting a sale. While recommender systems have become an essential computational component in major giant e-commerce companies, their complicated nature prevents SMB from readily adopting and understanding powerful Machine Learning (ML) algorithms and recommendation systems in their daily online operations. Therefore, the competitiveness of SMB in the US can significantly benefit from easier-to-use and affordable recommendation systems.

2 The value proposition

This project aims to reduce the ML development speed and complexity to accelerate the adoption of advanced recommendation systems by SMB, as pictured in Fig. 1, and help them thrive in the digital-first economy. Our team has developed a prototype to create recommendation systems [1,2] at a fraction of current time and cost.

3 The innovation

This project builds upon *our highly praised open-source system, AutoKeras* [3], which has become one of the most used AutoML systems (with over 8,000 stars and 1,300 forks on Github), to provide automated recommender systems. Our proposed framework learns the graph representation for any modal data to reveal the intrinsic relatedness of data at the feature or instance level. This project will allow SMB managers to automatically and efficiently develop their ML system.

4 The partnership

Our team has established a partnership with the following two SMB: Biarte offers various pantry products through sustainable end-to-end trade. This business is heavily transitioning its commercial operations to the web and will be at the center of this feasibility project as its transformation to an online presence is a sample of the challenges found in more SMB that our product will aim to solve. On the other hand, FreeFuse is a video-streaming platform with the mission to enhance idea-sharing by making learning content more customized and engaging.

The resulting prototype will enable FreeFuse and Biarte to share their user and product data to mutually benefit from a better recommendation engine.

5 Market analysis

Small and medium-sized businesses (SMB) worldwide are currently looking to improve their on-line operations to serve a market shifting to buy mainly on the web, which is often an overlooked collateral effect of the current lockdowns and stay-at-home restrictions worldwide. Recommendation systems are a valuable tool in e-commerce for SMBs because they reduce data overload by providing meaningful results to customers that increase the possibility of converting a sale.

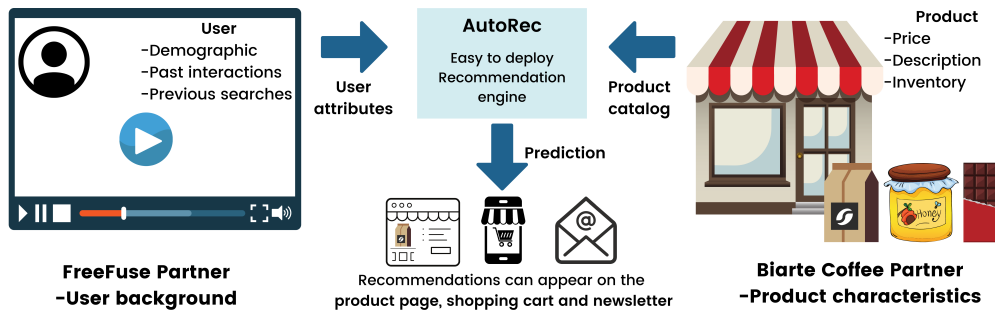


Figure 1: AutoRec application in e-commerce's recommendation systems.

Allowing SMBs to build and deploy robust recommendation systems becomes more economically significant when knowing that only in the US, more than 47% of jobs come from SMBs.

Particularly, IBISWorld analysts indicate that the business analytics software industry will generate 86.7 billion US dollars in revenue in 2021 in the US; as shown in Fig. 2, the e-commerce market represents 17% of the market with a reported 11.3% annual growth over the past five years [4]. This industry heavily focuses on large enterprises with the data requirements and budget sizes necessary to take full advantage of recommendation systems, leaving an unmet need for smaller firms. **However, IBISWorld's report indicates that smaller businesses will become more critical to this industry over the next five years as they adopt enterprise technologies that will eventually become the standard.** We see an opportunity to serve this market by developing easy-to-use ML tools with intuitive user interfaces tailored to retail and e-commerce managing teams with a basic programming background.

Business analysts from IBISWorld also indicate that online businesses are already heavy users of enterprise software. Online companies are exceptionally prepared to adopt enterprise software because they already heavily use information technology as a core part of their business. The e-commerce market segment will continue growing as online businesses displace traditional competitors. Altogether, online companies generated more than 4.0% of industry revenue in 2020 of the business analytics market [4].

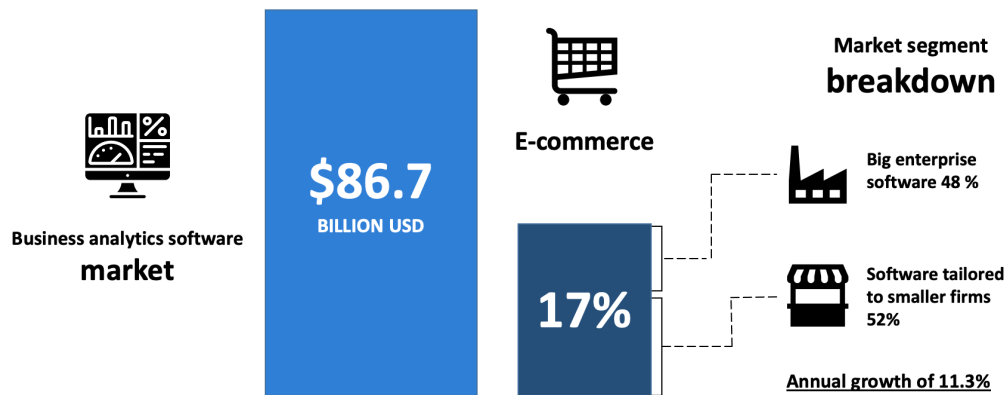


Figure 2: Business analytics software industry and e-commerce market segment concentration breakdown

6 Competitive technologies analysis

ML is a disrupting force across many industries; Statista projects ML-solutions' demand to grow more than 50% year to year over the next five years [5]. **Commercial recommendation technologies** such as Recombee [6], Crossing Minds [7], ExpertRec [8], and Strands [9], respectively, are increasingly investing in ML automation techniques. We are looking to complement these efforts

by further expanding our human-centered explainability technologies that are already visible in our published research and our popular open-sourced projects. **We firmly believe that our competitive advantage is our strong focus on ML automation, which we inherently combine with interpretable artificial intelligence (AI) to maximize human knowledge with intelligent machine support.** Specifically, our work in Auto-GNN can be used in interpretation because of the inherent user-item connection of GNN by knowing how two items are correlated, i.e., when a user has bought an item or shown a strong interest in a product, human experts can use the visual representation of such correlations to improve their recommendation systems.

6.1 Intellectual property

Our team has started early talks with the Office of Technology Transfer, but no relevant patent or technology disclosures has been made at the time of writing.

6.2 Timeline and milestones

PROJECT TIMELINE start November, 2021 - end August, 2022						Phase 1			Phase 2		Phase 3		Phase 4	
Research Objectives		Tasks		Team		m1	m2	m3	m4	m5	m6	m7	m8	m9
Milestones								M1		M2		M3		M4
O1: Developing Deep Recommender System	Learning Complex Interactions between User and Items			AC, XH, GS										
	Modeling Higher Order User-Item Relations			AC, GS										
O2: Automated Deep Recommender System	Automated CTR prediction			AC, GS										
	Automated Graph Neural Networks			AC, XH, GS										
	Recommender implementation and data collection in partner's systems			AC, XH, GS										
O3: End-to-End Robust Deep Recommender System	Automated Deep Recommender System: AutoRec			AC, GS										
	Evaluate Biarte and FreeFuse Data Response to End-toEnd system			AC, XH										
	Recommender implementation and data collection in partner's systems			AC, XH										
O4: R&D innovation advancement into a product offering	System Implementation and Result Evaluation of Experimental Setup			AC, GS										
	User Testing Results Action Items to Iterate R&D development			AC, XH, GS										
	Product-Market Fit Evaluation and PFI report			AC										

XH: Xia "Ben" Hu
AC: Alfredo Costilla-Reyes
GS: Graduate student

M1 Developing Deep Recommender System
M2 Automated Deep Recommender System for e-commerce using Biarte and FreeFuse data
M3 Crossplatform (Biarte and FreeFuse) End-to-End Robust Deep Recommender System
M4 R&D innovation advancement into a product offering

Figure 3: Recommendation system's MVP project timeline.

Overall, the development of a minimum viable product (MVP) will take 9 months. We will first start with implementing the system and then focus on enabling AutoML in our system. After that, we will spend half a year to polish and finalize our product based on our partner's feedback. **The Team** comprises expertise from Engineering Data Science and Entrepreneurship (Alfredo), Machine Learning and Artificial Intelligence (Hu), and Data Mining and Anomaly Detection (Graduate student). **PI Xia "Ben" Hu** is an Associate Professor at Rice University. He has published more than 100 papers in major data mining venues. His papers have received more than 7,000 citations and seven Best Paper Award (candidate), and he is the recipient of JP Morgan AI Faculty Award, Adobe Data Science Award, and NSF CAREER Award. **Dr. Alfredo Costilla Reyes** is a postdoctoral researcher at Rice University. He graduated from the Electrical Engineering doctoral program and Mays Business School from TAMU. Dr. Alfredo received Mexico's National Youth Award, for his contributions in technology and entrepreneurship.

7 Commercialization strategy

At this point, our business model will follow conventional practices in the business intelligence industry, divided into non-recurring and recurring revenue streams. **Non-recurring revenue streams** involve setup infrastructure costs such as AI & ML model building, platform installation, operator training, deployment expenses, and data preprocessing. Our **recurring revenue streams** may follow standard models of cloud container companies [10] such as AWS®. They include ML deployment operation, license fees, data volume analysis, and extra premium features related to security and data encryption.

References

- [1] Ting-Hsiang Wang, Xia Hu, Haifeng Jin, Qingquan Song, Xiaotian Han, and Zirui Liu. Autorec: An automated recommender system. In *Fourteenth ACM Conference on Recommender Systems*, pages 582–584, 2020.
- [2] DATA (Data Analytics at Texas AM) Lab. *AutoRec, an open-source automated machine learning (AutoML) platform extended from the TensorFlow ecosystem*, 2020 (Accessed February 9, 2021). <https://autorec.ai/>.
- [3] Haifeng Jin, Qingquan Song, and Xia Hu. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1946–1956, 2019.
- [4] Dan Cook. *IBISWorld US Industry (NAICS) Report 51121C. Business Analytics Enterprise Software Publishing in the US*, July 2020. Retrieved from IBISWorld database.
- [5] Tractica. *Revenues from the artificial intelligence software market worldwide from 2018 to 2025, by region (in billion U.S. dollars) [Graph]*, April 22, 2019 (Accessed October 6, 2020). <https://www.statista.com/statistics/721747/worldwide-artificial-intelligence-market-by-region/>.
- [6] PitchBook Data. *Recombee | Private Company Profile*, February 2021. Retrieved from PitchBook database.
- [7] PitchBook Data. *Crossing Minds | Private Company Profile*, February 2021. Retrieved from PitchBook database.
- [8] PitchBook Data. *ExpertRec | Private Company Profile*, February 2021. Retrieved from PitchBook database.
- [9] PitchBook Data. *Strands | Private Company Profile*, February 2021. Retrieved from PitchBook database.
- [10] Donna L Hoffman and Thomas P Novak. A conceptual framework for considering web-based business models and potential revenue streams. *International Journal of Marketing Education*, 1(1):7–34, 2005.