Artificial Intelligence Systems Engineering

System Goals

September 14, 2025

1 Introduction

Online Film Database (OFDb) is a website providing a range of features to users about films and their associated information, such as their casts, reviews and other news. In addition to dozens of existing features, OFDb is planning to implement a new system for premium users called RecommendR. Modelled after IMDb's "Top Picks" feature (IMDb Help Center, 2025), RecommendR will allow premium users to optimise their film selection experience by showing the top five up-to-date recommendations on their home page, and it will be powered by an in-house content-based recommendation system (Pazzani and Billsus, 2007). This document will define and evaluate several of the most important goals for this system in the context of OFDb's organisational objectives and leading indicators, as well as RecommendR's user outcomes and model details.

2 Organisational Objectives

The consideration to build RecommendR is principally motivated by OFDb's organisational objectives (Hulten, 2018); to increase revenue generated through advertising and premium account subscription fees by 5% and 10%, respectively. Both goals can be measured by calculating the total revenue generated per reporting period for each income stream and comparing to previous financial statements. While these are meaningful goals for OFDb as an organisation, they are not useful goals for the teams implementing RecommendR for three reasons. Firstly, advertising and premium user revenue are only loosely coupled to the performance of RecommendR, thus the system cannot be optimised based on these goals. Secondly, they are slow indicators with significant lead time between implementation and measurement. Thirdly, they are affected by many extrinsic factors, such as market forces (Hulten, 2018). Therefore, the correlation between these goals and a successful implementation of RecommendR is likely to be positive but weak.

3 Leading Indicators

Leading indicators are correlated with future success, and the two main indicators are user sentiment and user engagement (Hulten, 2018). OFDb plans to assess sentiment directly via a program of regular, digital surveys, and indirectly via A/B testing, designed to test new features against a baseline. On RecommendR's release, OFDb aims to achieve 90% positive sentiment across the user base for both sentiment measures. This measure can act as a proxy for how likely the users are to keep their subscription, thus relating to the broader organisational goal of increasing premium account revenue. User engagement gauges how much the customers use the product and can be measured by counting the average number of interactions with the website each user makes per day. OFDb expects a 5% increase in user engagement post RecommendR's release. Measuring user engagement can act as more than a passive indicator; it may directly contribute to the organisational objectives by being used as evidence to secure more lucrative advertising contracts, by proving that OFDb has an engaged and broad user base.

4 User Outcomes

User outcomes focus on setting goals pertaining to the user experience, and as such often relate strongly to model implementation decisions. Firstly, OFDb expects recommendations to be up to date, with a maximum wait of one day between breaking news and recommendation updates. Secondly, OFDb expects that, for each set of five film recommendations, at least one should be a film the user would watch next. It may be difficult to obtain high quality and consistent feedback from users on the performance of RecommendR. Therefore, investing in building a supervised test suite with examples covering a wide spectrum of user and recommendation combinations is advisable. As well as contributing to increasing premium account revenue, user sentiment and user engagement, a successful implementation of this user outcome links strongly to increasing advertising revenue. If OFDb can demonstrate accurate recommendation performance, then prospective advertisers will be more likely to trust the platform to recommend relevant advertisements to the users.

5 Model Details

Constructing goals around model details is valuable because they can be used to directly optimise the model. The model should be optimised to account for the different kinds of errors a user may experience: false positives and false negatives. OFDb has decided that, if necessary, it is acceptable to trade recall for precision. In the extreme case, supplying no recommendations will provide a worse user experience than providing poor recommendations; providing none may give the impression that RecommendR is broken. Therefore, the model will be optimised to have a precision@5 $\geq$ 20%, with the final metric chosen to align with the prior stated user outcome.

The RecommendR system must also be architected to be able to account for daily updates, such as new film announcements and other metadata changes, while maintaining the mandated precision@5 score. This will ensure a good user experience and further alignment with the user outcomes.

The average cost of recommendation inference per user, measured as the cost of compute and resources required to generate a recommendation averaged across users monthly, must be no greater than a fraction of the premium account subscription fee, to be determined by a future cash flow analysis. This will maximise any post-release profit margin increases, but may result in a more complex system architecture, potentially resulting in a larger up-front investment.

Model details goals do not capture the user's holistic system experience, because the goals focus on optimising aspects of low-level system architecture and performance. Therefore, considering these goals in partnership with leading indicators and user outcomes is critical to anticipating whether the system is likely to positively contribute towards the organisational objectives.

6 Goal Interrelationships

As has been highlighted throughout this document, the proposed goals are clearly related to each other, although the strength of the relationship varies significantly. They can be represented as an n-ary tree structure, starting from organisational objectives and ending at model details, with each layer offering goals that are more concrete and measurable than the last. Figure 1 shows a diagram inspired by fault tree analysis (Lee et al., 2009), where each system state block and primary failure event is the antithesis of a proposed goal. The diagram includes anticipated risks at multiple levels as possible failure events, representing the reality that goals can be put at risk by a multitude of internal and external factors.

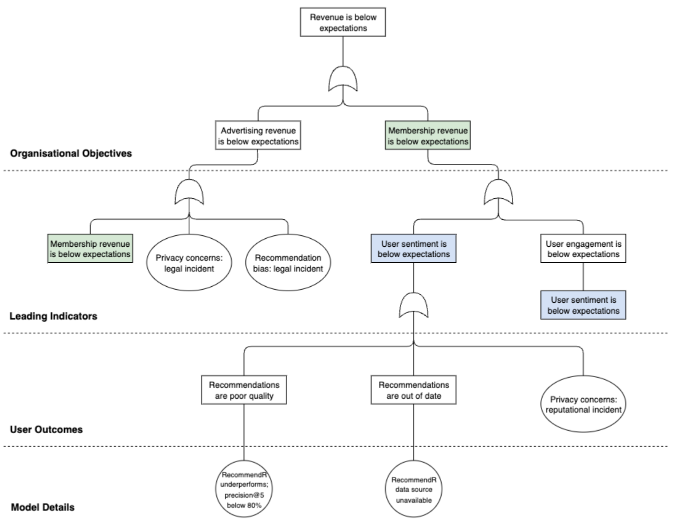

Figure 1: A modified fault tree analysis diagram, detailing the relationships between the goals and their anticipated risks.

7 Risk Analysis

The chance of achieving the aforementioned goals increases if accompanied by a risk analysis. Since this is OFDb's first AI-enabled system, there will be significant cost associated with data governance and management, infrastructure, and model engineering; a cash flow analysis must be conducted to evaluate if the system makes financial sense.

The stochastic nature of content-based recommendation systems may lead to inconsistent or surprising behaviour, which may harm the user experience in ways not fully captured by, for example, the model details goals. These technical risks may manifest themselves as the system making a recommendation that a user finds offensive. Data provides an additional technical risk vector; it is imperative that OFDb invests in securing a trusted and sustainable source for the data needed to power RecommendR, as well as the infrastructure and expertise needed to maintain it.

Processing user information in RecommendR introduces privacy concerns because of the potential leakage of protected characteristics and personal preferences, that may be implicitly learnt by the content-based recommendation system. This opens OFDb to additional legal and reputational risk if this data is exposed or mishandled, for example, by selling it to third parties.

OFDb is exposed to further legal risk by the system's inherent potential to exhibit bias. This may manifest as, for example, RecommendR tending to disproportionately recommend films with predominantly ethnic minority casts to users with non-western names. This may disadvantage certain films based on protected characteristics, likely an unfavourable and legally fraught proposition. Furthermore, film studios may figure out how to manipulate RecommendR to disproportionately recommend their films by engineering the film metadata, reducing the integrity of the system and exposing OFDb to further commercial risk. These risks can be managed by investing in a dedicated team to monitor holistic system integrity and behaviour, taking remedial action when appropriate.

8 Conclusion

Recommender systems are well established, so RecommendR is a relatively low risk endeavour in that precedent exists. However, careful consideration must still be given to the goals for the system in the context of OFDb's specific use case to ensure success against their organisational objectives. This document outlines several key goals and considers their justification, measurement and interoperation, finding that dependencies exist among all of them, despite often being only indirectly related. Lehtonen's framework (Kruhse-Lehtonen and Hofmann, 2020) shows there are many complicated and interrelated factors that contribute to the realisation of the vision and strategy for an AI-enabled system, so accounting for every factor when creating system goals is effectively impossible. Thus, even with the detailed and carefully considered goals as outlined in this document, success is not guaranteed.

References

Hulten, G., 2018. Defining the intelligent system's goals. In: *Building intelligent systems: a guide to machine learning engineering* [Online]. Berkeley, CA: Apress, pp.35–49.

Available from: https://doi.org/10.1007/978-1-4842-3432-7_4.

IMDb Help Center, 2025. *What to watch FAQ* [Online]. Available from: https://help.imdb.com/article/imdb/discover-watch/what-to-watch-faq/GPZ2RSPB3CPVL86Z#.

Kruhse-Lehtonen, U. and Hofmann, D., 2020. How to define and execute your data and AI strategy. *Harvard data science review*, 2(3).

Lee, W.-S., Grosh, D.L., Tillman, F.A., and Lie, C.H., 2009. Fault tree analysis, methods and applications a review. *Ieee transactions on reliability*, 34(3), pp.194–203.

Pazzani, M.J. and Billsus, D., 2007. Content-based recommendation systems. In: *The adaptive web: methods and strategies of web personalization*. Springer, pp.325–341.