A picture containing text, accessory, outdoor object

Description automatically generated

# Proposal Information

|  |  |
| --- | --- |
| **Track** | 🗹Talk – a 20-minute technical presentation, followed by up to 10 minutes of Q&A  ☐ Poster – a poster and short, recorded video. In-person posters will staff a booth at the Microsoft Conference Center in Redmond. Online posters will present a lightning talk and staff a virtual booth online. You do not have to create the poster or video as part of the submission process—please wait to do that until after the proposal is accepted.  ☐ Talk and a Poster – you can submit the same proposal for both Talk and Poster, and your proposal may be accepted for one, the other, or both.  ☐ Hands-On Lab – a hands-on lab that is taught in person  ☐ Tutorial – a 60-minute or 90-minute in-depth lecture with Q&A with the audience  ☐ Panel / Special Event  ☐ Social Activity  **To preserve a double-blind reviewing approach, please do not include your name(s) in this proposal.** |
| **Session Title (required)** | Active Learning Implementation: framework and lessons learned |
| **Session Objectives (required)** | In this talk, we will present a reference implementation of Active Learning using Azure technologies, MLOps methodology and share insights and lessons learned |
| **Session Audience**  **(required)** | *General audience familiar with machine learning concepts* |
| **Description (required -- proposals that do not include a description in this template will not be accepted)** | **Business objective:** Most supervised machine learning models require large amounts of data to produce good results. However, the process of labeling data can be time-consuming and expensive that represents a big barrier to machine learning development and adoption.  Active Learning is the methodology to prioritize the data which needs to be labelled in order to have highest impact to model training and a mechanism to run continuous training to deployment process. Active Learning allows ML model deployment to start with small amount of labelled data and low performance but can get better faster over time with lower labelling effort compared with investing in labeling large amount of data to start with.  **Insights and learnings**  In this talk, we will present a reference implementation of Active Learning using Azure technologies and MLOps methodology and share best practices in making Active Learning works.  1. Active Learning Implementation Framework: the detail design of the framework with components such as Model Monitoring, AML Labeling, Github action, AML, MLOps automation to select data, label, training and deployment. This is available in github as reusable frameworks at https://github.com/microsoft/MLOpsTemplate/tree/main/src/active\_learning\_cv  2. Best practices: comparing effectiveness of three different data prioritization: Least confidence, Margin sampling and Entropy. How to prepare training dataset for effective incremental learning. Other lessons learned in the training process.  **Technologies used**   * Azure ML * Azure ML Labeling * Github action * Azure Data Explorer * Azure ML computer vision AutoML   **Key takeaways**   * A reusable implementation & framework for your Active Learning project, including automation script, MLOps template, model, and data monitoring library. * Understanding effectiveness of different data prioritization techniques, best practices in forming training dataset and training process   *[Describe the objectives, content, and key takeaways of the session. Your description should make clear why the MLADS audience should care about your work—for example, why your solution is needed, how it compares to others (in and outside Microsoft), and the impact of your approach. If you can show customer value for the work, please do so.*  *See the Checklist below the submission form for guidelines for what to include to meet conference expectations around* ***Conference Focus****,* ***Analytical Depth****, and* ***Business Impact****.*    *Your proposal should include the following sections:*   1. *Business objective :* 2. *Insights and learnings that you will share -- including those related to how to think about building end-to-end scenarios that include AI* 3. *Production status : Customer value that the work is delivering* 4. *Technologies used* 5. *Key takeaways (success measurement / lessons learned). In particular, we are looking for how you can generalize what you learned from the work and make it relevant to the MLADS audience members, who will have a diverse set of backgrounds and experiences, and may not be well versed in your particular context.*   *You can include other sections as well.]*  *Note: You are welcome to submit supplementary materials, but make sure that your proposal is described in full in this form. For example, submitting an extremely brief submission form and attaching a 20-page research paper will make it difficult for reviewers to understand exactly what you intend to present at the conference.* |
| **Status of the work**  **(required)** | *[Describe the status of the work – MLADS focuses on work that is in production or close to it.]*  This framework is being implemented at two Microsoft’s customers.  ☐ In production since \_\_\_\_\_\_\_\_\_\_\_\_\_  🗹In production testing since Mar-2022\_\_\_\_\_\_\_\_\_\_\_\_\_  ☐ Validated offline  ☐ Completed A/B testing  ☐ Scheduled for A/B testing on \_\_\_\_\_\_\_\_\_\_\_\_\_  ☐ Other (please describe) |
| **Technologies**  **(required)** | *Active Learning strategies, Transformer-based NLP models, Named Entity Recognition, Text Summarization*   * Azure ML * Azure ML Labeling * Github action * Azure Data Explorer * Azure ML computer vision AutoML * MLOps * Active Learning |
| **Outline (optional)** |  |
| **Previous Presentations and Papers on this topic (optional)** | *[Has this material been presented elsewhere? And if you have previously presented on this topic at MLADS itself, explain what is new in this proposal, and why another session would be applicable at this time.]* |

# Checklist to Create a Strong Proposal

## þ Key Learnings

What insights will you share with the audience to help them improve their own skills and knowledge?

## þ Conference Focus – address conference focus in your description above

## A great proposal helps attendees:

## · Stay cutting-edge on the latest ML, AI, and data science tools and techniques; or

## · Understand best practices in machine learning, AI, and data science; or

## · Learn about Microsoft’s machine learning, AI, and visualization tools

* Adopt/adapt the learnings in their own context – that is, your proposal describes how your particular experience can be generalized to benefit those working in other contexts

And shows how learnings can be applied to solve real-world problems for Microsoft or its customers, improve customer experiences, or help make better business decisions

**Note:** The MLADS Conference primarily features applied work. Please discuss in your proposal the level of you work in terms of product maturity—has it been productionized, how did you evaluate success in production, what quantitative evidence do you have of customer value, do you have learnings from production / real users, and so on.

## þ Analytical Depth – address analytical depth in your description above

A great proposal is appropriate for a technical conference when it:

## · Contains core material that is strong in analytical / technical terms -- for example, if an ML model was applied, the proposal shows details around the model applied, train-test set and splits, data description and features, metrics defined, modeling results on the validation set, the alternative algorithms that were experimented with, and so on. For a more data science focused presentation, the talk should focus on data sources used, data cleansing that has been applied, and statistical analysis, as well as actionable insights.

* Describes the unique impact/contribution/novelty that the work is introducing
* Includes sufficient detail such that a skilled practitioner could reproduce the results
* Addresses principles and perspectives for [Responsible AI](https://microsoft.sharepoint.com/teams/Ethics/SitePages/AI-Ethics-Guidance-Directory.aspx) (fairness, transparency, reliability & safety, privacy & security, inclusiveness, and accountability), and includes discussion of how the project considers Responsible AI guidelines during design, development, and implementation
* Refers to related internal and/or external work and explains differences and/or improvements
* Includes appropriate visuals or tables that help the audience grasp the core ideas more easily

## þ Business Impact – address business impact in your description above A great proposal describes work that focuses on a good target and delivers demonstrated customer value. A good target:

## · Engages top management commitment and creates momentum

## · Focuses on generating insight rather than merely information

## · Is both ambitious and approachable—ambitious in that it has business impact, and approachable in that it has access to the resources and capabilities to succeed

## · Is focusing the analytical investment on Microsoft’s distinctive technologies, which serve our customers in a way differentiated from our competitors

Use the following table to assess the audience level of your proposal (100-level, 200-level, 300-level, 400-level):

# Engineering Sessions Proficiency Levels

|  |  |  |
| --- | --- | --- |
| Expertise level | Description | Example learning objectives for an engineering audience |
| **Foundational**  Synomyns: level 100, beginner, novice, introductory, overview, survey | This session is for a participant that is **new to this topic or has limited applied experience** and seeks to better understand the fundamentals. Covers topic concepts, functions, features, and benefits. | 1. ***Describe*** *the purpose, function, and implementation of a given theory, method, or function.* 2. ***Identify*** *typical real-world scenarios for a given theory, method, or function.* 3. ***Implement*** *a “hello world”/basic example of a given theory, method, or function.* 4. ***Identify*** *where to go for further training.* |
| **Intermediate**  Synomyns: level 200, intermediate, application, doing | This session is for a participant with a **general understanding of key principles or some applied experience** and is seeking an increased depth of knowledge on this topic. Covers core approaches and skill development in real-world scenarios. | 1. ***Explain*** *theoretical or functional architecture, scenario integration and configuration* 2. ***Discuss*** *first and third party case examples* 3. ***Implement*** *a real-world functional application, visualization, or algorithm* 4. ***Explain*** *edge cases, corner cases, and common pitfalls* |
| **Advanced**  Synomyns: level 300, advanced, synthesis, PBL, how, why, impact | This session is for a participant with a **solid understanding of key principles and applied experience** and is seeking an increased depth of knowledge on a particular aspect of this topic. Covers the topic and subtopics in detail with problem-solving real-world scenarios. | 1. ***Improve*** *upon real-world implementations by audience member* 2. ***Diagram*** *architectural constraints/considerations of the relevant implementation* 3. ***Solve*** *corner and edge cases* 4. ***Demonstrate*** *code work-arounds and troubleshooting techniques* 5. ***Articulate*** *implementation constraints and considerations in a whitepaper* |
| **Expert**  Synomyns: level 400, expert, mastery, SME, evaluation, creation, innovation, strategy | This session is for a participant who is an **acknowledged thought leader in the use and implementation of the topic or skill or is seen as a resource to teach and coach others on the topic** and is seeking to engage in the latest developments. Typical sessions are smaller and more specific in nature. Provides expert-to-expert interaction and coverage of specialized topics. | 1. ***Demonstrate*** *efficiency skill in functional architecture, scenarios integration, or algorithms* 2. ***Deliver*** *design concepts, security and implementation approaches and tradeoffs* 3. ***Demonstrate*** *strategic and business impacts of the technology* 4. ***Articulate*** *theoretical models of expansion* 5. ***Design*** *future-state algorithms or models* 6. ***Advise*** *on standards creation and regulation compliance* |