IMRating	CommentsIntelMerit	ROIRating	CommentsROI	BroaderImpactRating	CommentsBroaderImpact
22-049					
4	First I should state this is not my field. This work will seek the development of an advanced Machine Learning tool specific to the needs of scientific experiments which generate very large data sets. The work will seek to combine in a new and innovative way currently existing analysis tools and allow integration of new tools as they become available. The machine learning platform will rapidly optimize the selection of the set of tools to be used for the needs of the project. The process of rapidly completing the selection of the analysis set for individual projects will require very innovative application of the ML skills which the proposing team has.	4	I feel the potential is good but this is now a crowded field.	4	This work will provide new capabilities enabling BNL to do our research work more efficiently helping to demonstrate BNL's leadership in data-intensive research.
5	The team are have shown a clear understanding of the type of problems that can be solved with AI and machine learning (ML) approaches. Their current data examples (ATLAS and DUNE) are robust and well suited to their desired model training methods. In addition, their stated intent to make their software tools inside of a portable frameworks means that said tools will like be useful to many different projects around the lab(s), outside of ATLAS and DUNE.	4	The team has focused strongly on portability and scalability. If successful, this could allow their tools to be deployed for a wide range of projects that focus on large-scale, data-intensive, scientific applications. This could mean a very large return on investment if their tools become widely used at facilities like particle accelerators.	5	Given the team's focus on portability, I would imagine that their success with analysis of large data sets (ATLAS and DUNE), would allow them to easily alter the algorithms such that they can be trained at used with other experiments that need highly efficient analysis using an application agnostic ML service environment. You can imaging these services being very useful at synchrotrons such as ALS or NSLS-II.

IMRating	CommentsIntelMerit	ROIRating	CommentsROI	BroaderImpactRating	CommentsBroaderImpact
4	"FaaS may not be a well suited for data intensive service as the communication cost is usually very expensive. GAN training took 100 GPU days could be true but GAN training need more fine tuning and improving, in general due to the nature of min-max optimization. Portabilities for deep learning is also interests on the industries and the industry made progresses such as XLA (accelerated linear algebra) compiler framework, which already working for GPU, TPU and some AI accelerators. Given the limited resources, maybe focus more on easy to use scalability aspect would make higher value, which has enough research topics.Another important aspect is accounting and most open computing platform restrict free access to the limited capabilities unless the user pays it. So, user authentification needs to be coupled to accounting system and provide a mechanism to communicate the end user would be the practical system although investigating this would not be priority in the prototype system implementation. ML workflow is data intensive and there are many existing ML workflow. Thus the backend design efforts may need to embrace how ML framework communities are working, not necessarily relying on the previous success.Also please don't forget to what industries are doing such as MLFlow (MLflow - A platform for the machine learning lifecycle | MLflow), which is already doing significantly in the industries although it is not optimized for HPC system. And there are some commercial products as well (from data preprocessing to hyperparameter optimization, so that it reports final results with the summary of analysis but need to pay license fees).Some of open source framework also integrated Jupyter with HPC backends and it might be worth to extend such framework instead of starting from scratch."	4	Due to the user facilities, our lab have a high chance to get RoI.	4	The community is also moving toward the proposed direction and thus if we make a breakthrough, then it will make an impact.

IMRating	CommentsIntelMerit	ROIRating	CommentsROI	BroaderImpactRating	CommentsBroaderImpact
3	The proposed work is extremely important for BNL, however it is primarily an engineering effort with on going maintenance costs for the infrastructure. While the plan is feasible, the estimated budget is at least 4 x too small.Also, the few details of how this proposal solves challenges that prevent use of AWS, Azure, etc.. are not clear. Other than providing the infrastructure (which is important and should be funded by BNL) , the proposal did not have a detailed path for how the proposed work will enable 1) AI enhanced detectors, accelerators, and sensors; 2) optimal experimental design and steering specified in the LDRD call. If the authors could enumerate half a dozen or so specific applications and why their ML training algorithms need such a large infrastructure, that would go a long way in justifying the proposal.The authors also mentioned fast calo gan as one justification for this effort. GAN's are notoriously hard to train and substantial effort is going into developing Autoencoder based alternatives. Placing the justification on this sole use case given the pace at which ML is developing would be ill advised.	3	None	4	There is potential for broader impact on the lab, however there needs to be a clearer vision driven by a more diverse sampling of specific use cases.
Avg 4		3.75		4.25	