AI Explainability WG

October 2024

September 2024 updates

- TrustyAI core / service
  - Current: 0.21.0 release
    - https://github.com/trustyai-explainability/trustyai-explainability/releases/tag/v0.21.0
    - quay.io/trustyai/trustyai-service:v0.21.0
- TrustyAI operator
  - Current: 1.26.0 release
    - https://github.com/trustyai-explainability/trustyai-service-operator/releases/tag/v1.27.0
  - quay.io/trustyai/trustyai-service-operator:v1.27.0

What's new?

TrustyAI - What's new?

- **TrustyAI core/service 0.21.0**
  - RHOAIENG-13242: Update io-grpc from 1.59 to 1.68
  - RHOAIENG-13242: Override the version of protobuf-java in the service until io-grpc moves from 3.25.3 to 3.25.5
  - Integrate Tensor Type into service consumers and data types by @RobGeada in #622
  - RHOAIENG-13625: Add errors to termination-log
  - Enhancement: ImageDrift Generalization
  - RHOAIENG-10826: Truncate drift request output to relevant fields
  - Add unit test for /metrics endpoint

TrustyAI - What's new?

- **TrustyAI operator 1.27.0**
  - RHOAIENG-13625: Add DBAvailable status to CR

Current work

TrustyAI - LM-Eval

- Work taking place in dev branch
  https://github.com/trustyai-explainability/trustyai-service-operator/tree/dev/lm-eval
- Target release ODH 2.20
- Support for custom datasets
- Support for custom unitxt cards/recipes
- Initial support for Kueue
- Replaced Job pushing status to operator pulling status
- Removed hardcoded user id
- Merge into main branch will happen 21st October 2024

Roadmap

TrustyAI 2024 roadmap

- ○ KServe explainer integration
- ○ Detoxification fine-tuning
- ○ Saliency Explainers
- ○ Guardrails
- ○ LM-Eval

Other topics