

Knowledge Discovery Interest Group

Raffaele D'Abrusco

CENTER FOR

ASTROPHYSICS

HARVARD & SMITHSONIAN

KD-IG session@2021FallInterOp

- ⊕ Three interesting talks about relevant KD scientific applications and methodologies
 - ⊙ A. Mahabal: *Data Sheets and Model Cards*

An approach to standardized description of datasets and models that can be used as benchmarks for KD applications in astronomy and can favor data fusion.
 - ⊙ P. Skoda: *SDSS redshift prediction based on Bayesian Deep Learning*

Application of Bayesian deep learning technique which provides uncertainty estimations to the determination of spectroscopic redshifts
 - ⊙ R. Martinez-Galarza: *Harvesting outliers: data barriers to turn anomalies into discoveries*

Selection of anomalous objects from large astronomical datasets and the problem of interpretability when selection occurs in non-physical feature spaces
- ⊕ Open discussion regarding KD-IG priorities

Reinforce priorities

- Making sure that data mining algorithms can run on non-tabular data types
 - ⊕ Polymorphic data access
 - ⊕ Support DM-based labeling of data
- Handling uncertainties and probabilistic measurements (pdf)

Collect new ideas

- Stay abreast of new developments in KD-relevant methods
 - ⊕ Solicit a robust range of contributed talks for KD-IG sessions
 - ⊕ Invite members of interesting projects to speak at KD-IG sessions
 - ⊕ Planning for a future IVOA InterOp plenary session with significant intra-IVOA and community involvement

Set goals and *modus operandi*

- Document interests, needs and recommendations
 - ⊕ Regular, single-topic focused, “in-between InterOps” meetings
 - ⊕ Drafting an *implementation note*
 - ⊕ Document one or a few related recommendations in a *note*

<Insert classical Uncle Sam's
"We want you" image>

If you are interested in being the V Chair for KD, please send a note to the IVOA Executive committee (or contact me if you want to know more before).