

ORES Preparation II

Tom Gülenman

14. Dezember 2018

Disclaimer: No guarantee for the correctness of information / explanations / sources is given.

1 Basic API Usage

- <http://ores.wmflabs.org/v3/scores/enwiki/?models=draftquality|wp10&revids=34854345|485104318>

Scoring, Context = enwiki, Models (2), RevIDs (2)

2 ORES Scoring Documentation

<https://ores.wikimedia.org/v3/#/>

- <https://ores.wikimedia.org/v3/scores>

Lists all available contexts (wikis) for scoring

```

"enwiki": {
  "models": {
    "articlequality": {
      "version": "0.6.1"
    },
    "damaging": {
      "version": "0.4.0"
    },
    "draftquality": {
      "version": "0.1.0"
    },
    "drafttopic": {
      "version": "0.1.0"
    },
    "goodfaith": {
      "version": "0.4.0"
    },
    "wp10": {
      "version": "0.6.1"
    }
  }
},
"enwiktionary": {
  "models": {
    "reverted": {
      "version": "0.4.0"
    }
  }
},
"eswiki": { ...

```

Models: reverted, damaging, goodfaith, articlequality, draftquality, drafttopic, wp10, pagelevel, itemquality

- <https://ores.wikimedia.org/v3/scores/{context}>

For example **enwiki**

Shows the specified context only

- <https://ores.wikimedia.org/v3/scores/{context}/{revid}>

Shows all scoring information on an edit of the specified context

Find revision ID of any article: top right: View history → click on time/data-link → ID displayed in URL at the end

```

{
  "enwiki": {
    "models": {
      "articlequality": {
        "version": "0.6.1"
      },
      "damaging": {
        "version": "0.4.0"
      },
      "draftquality": {
        "version": "0.1.0"
      },
      "drafttopic": {
        "version": "0.1.0"
      },
      "goodfaith": {
        "version": "0.4.0"
      },
      "wp10": {
        "version": "0.6.1"
      }
    },
    "scores": {
      "870259881": {
        "articlequality": {
          "score": {
            "prediction": "B",
            "probability": {
              "B": 0.5301044828795669,
              "C": 0.09410444935893472,
              "FA": 0.10609462640332888,
              "GA": 0.2630374655004465,
              "Start": 0.005353767142851241,
              "Stub": 0.001305208714871623
            }
          }
        }
      },
      "damaging": { . . .

```

- <https://ores.wikimedia.org/v3/scores/{context}/{revid}/{model}>

Furthermore a model can be specified

- Add `/?model_info={model_info}` to any query (at the end!)

For example: `model_info = type, statistics`

```
{
  "arwiki": {
    "models": {
      "damaging": {
        "type": "GradientBoosting"
      },
      "goodfaith": {
        "type": "GradientBoosting"
      }
    }
  }, ...
}
```

3 Specifying parameters

- From the paper: https://ores.wikimedia.org/v3/scores/enwiki/?models=damaging&model_info=statistics.thresholds.true.'maximumfilter_rate@recall>=0.75'

(spaces: maximum filter_rate @ recall >= 0.75')

Information:

- only **damaging**-model
- model.info: **statistics** ← this is what we need for parameters in our *Visualisierung des Systemverhaltens*

4 Parameters

4.1 Model: Damaging

4.1.1 Thresholds

Interesting: model.info=statistics.thresholds.true

Show all true values of parameters (that are not in red below) with the threshold value from 0.002 to 0.987 (not every step, but many!)

Without “.true” → we got double the results divided in false: [...] and true: [...]

4.1.2 Different parameters

- !f1
- !precision
- !recall
- accuracy
- counts
- f1
- filter_rate
- fpr
- match_rate
- pr_auc
- precision
- rates
- recall
- roc_auc

4.2 Model: Articlequality

4.2.1 Thresholds

Interesting: model.info==statistics.thresholds

→ “.true” doesn’t work

Lists for each article quality (Stub to FA) the values for each parameter, given a threshold

```

{
  "enwiki": {
    "models": {
      "articlequality": {
        "statistics": {
          "thresholds": {
            "B": [
              {
                "!f1": null,
                "!precision": null,
                "!recall": 0.0,
                "accuracy": 0.035,
                "f1": 0.067,
                "filter_rate": 0.0,
                "fpr": 1.0,
                "match_rate": 1.0,
                "precision": 0.035,
                "recall": 1.0,
                "threshold": 0.002
              },
              {
                "!f1": 0.003,
                "!precision": 1.0,
                "!recall": 0.002,
                "accuracy": 0.036,
                "f1": 0.067,
                "filter_rate": 0.002,
                "fpr": 0.998,
                "match_rate": 0.998,
                "precision": 0.035,
                "recall": 1.0,
                "threshold": 0.003
              }
            ]
          }
        }
      }
    }
  }
}

```

4.2.2 Different parameters

Same as above.

5 Clearing up confusion about parameters from before

Paper page 15: “when a **threshold** is set on 0.299 likelihood of damaging=true, then you can expect to get a **recall** of 0.751, **precision** of 0.215, and a **filter-rate** of 0.88”

- **Threshold** on 0.299 of damaging=true → if probability that an article is damaging ≥ 0.299 , then it is declared as damaging
- **Recall**: 0.751 → given threshold catches 75.1% of (most) damaging edits
- **Precision**: 0.251 → approx. $\frac{1}{4}$ of edits, declared as damaging, actually are

- **Filter-rate**: 0.88 → work of vandal-fighters reduced by 88%

6 Specifying parameters II

6.1 Querying formulas

Paper:

- https://ores.wikimedia.org/v3/scores/enwiki/?models=damaging&model_info=statistics.thresholds.true.'maximumfilter_rate@recall>=0.75' → write as "**maximum filter_rate @ recall >= 0.75**"
- **maximum recall @ precision >= 0.9** produces:

```
{
  "enwiki": {
    "models": {
      "damaging": {
        "statistics": {
          "thresholds": {
            "true": [
              {
                "!f1": 0.984,
                "!precision": 0.969,
                "!recall": 1.0,
                "accuracy": 0.968,
                "f1": 0.154,
                "filter_rate": 0.997,
                "fpr": 0.0,
                "match_rate": 0.003,
                "precision": 0.902,
                "recall": 0.084,
                "threshold": 0.916
              }
            ]
          }
        }
      }
    }
  }
}
```

- **maximum** can be replaced with **minimum**
- **>=** can be replaced with **<=**
- Both parameters can be replaced with any from the list above (except threshold ofc, which is not a parameter)

Parameters to use in ORES' API queries

Example format: `https://ores.wikimedia.org/v3/scores/enwiki/?models=damaging&model_info=statistics.thresholds.true.'maximumfilter_rate@recall>=0.75'`

(spaces: `maximumfilter_rate@recall>=0.75'` or `"%20"` in HTTP)

- `!f1`
- `!precision`
- `!recall`
- `accuracy`
- `f1`
- `filter_rate`
- `fpr`
- `match_rate`
- `precision`
- `recall`