*Metrics Research*

Before deciding on how to measure our system, we researched standards for waste sorting that we could compare Recyclotron to. We wanted to see what was a realistic example of success for recycling and sorting.

The British government's 2018 Resources and Waste Strategy [1] helped us compare metrics, with their strategic indicators helping us model success better.

Unfortunately the metrics focused on measuring national recycling performance, rather than a single location's performance.

We also discovered similar smart-bin projects such as Trashé [2] and SortBot [3]. Only Trashé had performance we could compare with (precision, recall & mAP), as SortBot was not fully integrated with Machine Learning.

We found a BBC article stating the sorting performance of three English recycling centres. They both successfully recycle around 85\% of recyclables they receive.[4] The other 15\% of recyclable material is lost.[4] This informs us that we should be classifying no more than 15\% of our recyclables as trash.

Finally we found a case study comparing various sorting standards in England, Europe and North America.

It gave us a measure of how pure recycled categories need to be, before they can be resold on the market as new materials.

Generally, UK Markets forbid the resell of materials that are more than 1\% contaminated with other materials.[5]

For the French Market, Eco-Emballage specifications gave targets for precision in various categories of rubbish.[5]

Now with some indicators of performance to aim for, we had to decide how best to measure them.

Entropy was a possible metric, as it can be used to describe the purity (or order) of a set. It can be easily abstracted and scaled with additional categories (or sets). However, Entropy accounts for set size, giving impurity as a number, with 0 being a pure set. It is also symmetric, with a 90\% accurate classifier scoring the same as a 10\% accurate classifier.

All our targets relate to precision in each recycling category. This justifies the use of precision as a test metric. However we also want to retrieve as many recyclable items as possible to minimise landfill waste, so we need a metric that considers recall but can still be interpreted as precision.

As a compromise, we're using mean Average Precision (or mAP) as our overall metric. This is because mAP's calculation incorporates precision and recall for every category.

Average precision is a commonly used in information retrieval contexts, which is similar to our 'recyclable retrieval' use case. AP is defined as the average precision across different thresholds for the predicted probabilities. AP is computed as area under the precision-recall curve. mAP is simply the mean AP across categories.

Below is a table of performance (or success) indicators, what they tell us, what real they're comparable to, how we measure them (metric) and what a success is (is it on target?).

Total Truly Recycled is our overall measure of performance, which uses mAP. Level of Trash Contamination indicates how much trash is misclassified. This is helpful in showing us false positives, which we want to eliminate to reduce recycling centre costs (as per report 2).

The final 5 performance indicators show us how well recyclotron performs in specific categories.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Performance Indicator* | *Tells us...* | *Why we need to know* | *Metric* | *Desired Target* |
| Total Truly Recycled  (Overall Performance) | How much recyclable material we successfully sort as recyclable | Comparable to Trashé and the Kirklees, Hull & Greenwich recycling centres performance. | mAP | >=85% |
| Overall Trash Contamination | How much trash falls into other categories | Comparable to UK Market standards of contamination | 1 - Trash Recall | <= 1% |
| Metal Purity | How much category consists of metal | Comparable to French Market standards of recycling purity | Precision of Metal Category | >=45% |
| Paper Purity | How much category consists of paper | Comparable to French Market standards of recycling purity | Precision of Paper Category | >50% |
| Cardboard Purity | How much category consists of cardboard | Comparable to French Market standards of recycling purity | Precision of Cardboard Category | ~95% |
| Plastic Purity | How much category consists of plastic | Comparable to French Market standards of recycling purity | Precision of Plastic Category | >90% |
| Glass Purity | How much category consists of glass | Comparable to French Market standards of recycling purity | Precision of Glass Category | 100% |

Sources on metrics:

[1] [Our Waste, Our Resources: A Strategy for England](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/765914/resources-waste-strategy-dec-2018.pdf?fbclid=IwAR2HzCumTjjwpC5eZMCclj7S1b_U5koHAgdt_ld76MsZinnsp1GBMqhZ7ak)

[2] [Trashe Smartbin](https://devpost.com/software/trashe?fbclid=IwAR07fHWuiQxyECnnLv381zXNnA4Ir5sauMqHf5aZRwhWCBHa5AKoL9vVmKk)

[3] [SortBot: Self-sorting Recycling Machine](http://www.ee.ic.ac.uk/guo.liew15/yr2proj/Group%2006_Self%20Sorting%20Recycling%20Machine_Cong.pdf)  
[4] [Rejected recyclable waste up 84% in England since 2011, data shows](https://www.bbc.co.uk/news/uk-37159581)

[5] [An introduction to MRFs and comparison of sorting operations](http://www.wrap.org.uk/sites/files/wrap/MRF_v6_19Dec06_LC.pdf)

*Unused:*

[Measuring recycling performance - Commercial Waste](https://cleanstreets.westminster.gov.uk/measuring-recycling-performance/)

[Input Data Report](https://cleanstreets.westminster.gov.uk//wp-content/uploads/2015/09/WCC-MRF-performance-Q4-2014.pdf)

[Data expert ranks recycling metrics](https://resource-recycling.com/recycling/2018/06/19/data-expert-ranks-recycling-metrics/)

[Noting what's not recovered](https://resource-recycling.com/recycling/2016/10/10/noting-whats-not-recovered/)

[Start at the cart](https://resource-recycling.com/recycling/2018/04/02/start-at-the-cart/)

[What is the best validation metric for multi-class classification?](https://sebastianraschka.com/faq/docs/multiclass-metric.html?fbclid=IwAR1pj1r68XZWPYk5ySY_YOFIlzAx-x0QyeH5_w1ripKRmPclWdY5dXiSAmA)

[What is Entropy?](https://learn-eu-central-1-prod-fleet01-xythos.s3-eu-central-1.amazonaws.com/5d1b15b77a8ac/2372175?response-content-disposition=inline%3B%20filename%2A%3DUTF-8%27%27dt.pdf&response-content-type=application%2Fpdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20200308T161114Z&X-Amz-SignedHeaders=host&X-Amz-Expires=21600&X-Amz-Credential=AKIAZH6WM4PLYI3L4QWN%2F20200308%2Feu-central-1%2Fs3%2Faws4_request&X-Amz-Signature=6c31f1f368b1ab58a896afb6b9eb22aae3a07968fdc452a93bc9cf059d3c6454)

[Why is entropy useful?](http://www.inf.ed.ac.uk/teaching/courses/fnlp/lectures/04_slides.pdf)