

Machine Learning for Email

This page is to brainstorm and coordinate development for better handling email within Tiki using [machine learning](#). This is a great use case, and will serve as an example for integrating machine learning all over Tiki. The Machine Learning is with [Rubix ML](#) and the webmail is with [Cypht](#). See also [Email as a first-class citizen](#)

Below are some examples. Once we develop the code, they will be easily adaptable to your similar use cases with your data.

Classification use case

- Emails relevant to bookkeeping (invoices, statements, etc)
- Emails relevant to servers (server down, monthly uptime report)
- SPAM or irrelevant emails (Viagra, watch, loto). While mailbox has tools some always pass.
- Emails that can wait (news, magazines, etc. Nice to read but not essentials to productive work)
- Email that should be associated to a task, project or client
 - As per [Email as a first-class citizen](#): "Instead of one gigantic mail store, we should have a number of smaller ones that make sense to one's workflow (ex.: around projects, tasks, clients, etc.) and that can easily be shared and prioritized."

Potential workflow

1. Look at all these folders to train data
 - Past data
2. When new mail arrives here, here and here
 - If confidence level high, move to relevant folder
 - If confidence level is not high, propose to user and learn from answers

What we want to know

- Is this an email which requires action (ex.: a bill to pay or a server to tend to) or it's just information that can be analyzed one day if needed.

Inspiration

[NextCloud](#) uses [Rubix ML](#) as well:

- <https://github.com/nextcloud/mail/blob/master/lib/Service/Classification/ImportanceClassifier.php>
- <https://nextcloud.com/blog/nextcloud-mail-introduces-machine-learning-for-priority-inbox/>

Pablo Duboue

<https://sourceforge.net/projects/listosoreader/>

Related

- [Machine Learning for SEO and SEM](#)
- [Auto-Classification](#)
- [Naive Bayes classifier](#)