

What happens when click on “Mark As” or “Report As”?

PROBLEM

What happens when the **Admin** clicks on “Mark As” or “Report As”?

SOLUTION

Handling **false positives (FP) and false negatives (FN)** is a natural part of operating any email security system. Even the most advanced filtering engines may wrongly classify messages sometimes, especially during the initial learning phase or based on individual user preference. Newsletters, automated notifications, and marketing emails are common examples where the spam perception can vary from one recipient to another. For this reason, an effective system must not only detect errors but also provide fast and reliable tools to correct them and continuously improve accuracy.

The primary mechanism for managing false positives and false negatives is Machine Learning and the AI Classifier. The system continuously learns from mail flow, message classification, and user behavior, adapting its decisions over time. Administrators play a key role in accelerating and refining this learning process through two main actions available in the interface: **“Mark as”** and **“Report as.”**

When an administrator clicks **“Mark as”**, the action is used exclusively for **local Machine Learning training**. This tells the system how that specific message should have been classified and immediately improves the behavior of the Machine Learning engines on the appliance. No email sample is sent outside the system, and the correction applies only to the local environment. This option is particularly useful to speed up initial training after deployment or to quickly correct recurring misclassifications that are specific to the organization.

When an administrator clicks **“Report as”**, the action performs three operations at the same time. First, it trains the local Machine Learning engines in the same way as the “Mark as” action. In addition, it **sends the full original email sample**, along with the gateway’s analysis, to the Libraesva EsvaLabs.

The samples are securely transmitted to the Libraesva EsvaLabs through an encrypted communication and are stored offline and locally to the EsvaLabs with an AES256 encryption.

Security analysts at EsvaLabs review the messages, identify indicators of compromise, correlate them with other samples, and implement corrective actions to global reputation data and detection engines. They also update the privacy-first AI engine on your appliance on a daily basis in order to continuously improve it's capabilities. These frequent updates prevent the recurrence of the same false positive or false negative in the future.