AutolearningNotWorking

Why isn't autolearning working for me? (aka: "autolearn=no")

Lots of people seem to be confused by the "autolearn=no" statement in the default X-Spam-Status header. There are usually questions regarding whether or not "no" means SpamAssassin is not autolearning at all. What it actually means is that the specific message which includes the "autolearn=no" part was not autolearned, not that autolearning is disabled or somehow broken.

The three values that can be displayed are "no" (autolearning did not occur), "ham" (the message was learned as ham), and "spam" (the message was learned as spam).

If a message has already been learned by SpamAssassin, then that message will not be learned again. Therefore, if you run a message through SpamAssassin to see why it was classified as spam or ham, and it has already been learned, you will always get the result "autolearn=no". (To see this more clearly, use the "-D" flag, and you will see debug output explaining that the message has already been learned.)

Furthermore, the score used to trigger autolearning is somewhat different than the one reported in the final score; therefore a score displayed in the headers that obstensibly should trigger autolearning will not do so. Again, use the "-D" flag to SpamAssassin, and you will see the score that is used to determine whether or not autolearning will be triggered.

Finally, SpamAssassin requires at least 3 points from the header and 3 points from the body, to auto-learn as spam. If either section contributes fewer points, the message will not be auto-learned.

For more information, please read the Mail::SpamAssassin::Conf documentation,

Possible autolearn states

In SpamAssassin 2.5 and 2.6, there were only three states for the autolearn result:

- ham: the message was learned as ham (non-spam)
- spam: the message was learned as spam
- no: the message was not learned

In SpamAssassin 3.0, the result was enhanced to have six states:

- ham: the message was learned as ham (non-spam)
- spam: the message was learned as spam
- no: the specific message didn't achieve the proper threshold values and requirements to be learned
- disabled: the configuration specifies bayes_auto_learn 0 or use_bayes 0 and so no autolearning is attempted
- failed: autolearning was attempted, but couldn't complete. This happens if SpamAssassin can't gain a lock on the Bayes database files, etc.
- unavailable: autolearning not completed for any reason not covered above. It could be the message was already learned.