

BayesNotWorking

Bayes doesn't seem to be working for me!

There's a number of possible reasons this may be happening.

- if you're running SA 2.5x, the 'middle ground' of Bayes scores – around the 0.5 mark, where the classifier didn't find enough info in the message to make a reliable determination – does not produce a rule hit. In other words, you won't see BAYES_40 or BAYES_50 listed in the rules-hit report. This is normal. SA 2.60, however, has tiny scores assigned for these rules, since this was becoming an FAQ. 😞
- SA 2.60 will only use Bayes if the DB_File perl module is installed. The package itself will install fine, but the Bayesian filtering parts will not be used (since SA can be run just fine without Bayes). Run "spamassassin -D --lint" and keep an eye out for a line like:

```
debug: bayes: Not available for scanning, DB_File not installed
```

- you might be using spamd, and running 'sa-learn' as a different user. Check your /var/log/maillog syslog file, then the spamd switches and configuration, to see what username the mail is being scanned as. Obviously, if you learn messages into one user's database, the learning will not be used when scanning messages as a different user.
- there's a minimum threshold on how many messages must be in the Bayes database, before SA will use it while scanning. By default, there must be 200 ham messages and 200 spam messages learned before it will be used.

This is a default, conservative setting; SA's Bayes is often used in an administration-free mode, and we determined experimentally that lower thresholds require more 'babysitting'.

However, if you're happy to keep an eye on things, go ahead and reduce these numbers by using the 'bayes_min_ham_num' and 'bayes_min_spam_num' settings in your configuration file. Numbers as low as 10 will work fine, as long as you expect the odd mistake during the first few messages, and will run sa-learn to correct their classification. (This is called 'mistake-based training'.)

This can be checked by running "spamassassin -D --lint" and keeping an eye out for this line:

```
debug: bayes: Not available for scanning, only 0 spam(s) in
Bayes DB < 200
```

Still No Bayes Filtering? Watch out for the "rebuild" option!

If you find that SA never seems to learn messages, try using `sa-learn --dump magic` to find out more about your database. The line "nham" is the number of ham messages SA has learned, and the line "nspam" is the number of spam messages SA has learned.

There's a subtle gotcha in the options. The command `sa-learn --no-rebuild --spam spam_directory` works. However, the command `sa-learn --rebuild --spam spam_directory` will *not* work; it will rebuild but *not* learn!

To defer rebuilding until after all the messages have been learned, use:

```
sa-learn --ham --no-rebuild ham_directory
sa-learn --spam --no-rebuild spam_directory
sa-learn --rebuild
```