LearningMarkedUpMessages

Will the learner break if I learn SpamAssassin-modified spam messages?

No. If you run sa-learn on a mail folder that includes messages with the SpamAssassin markup (X-Spam-Status headers, "real" message encapsulated as a message/rfc822 MIME part, Subject header tagged etc.), it'll automatically remove the markup on the fly.

Note: sa-learn does *not* extract RFC-822 attachments if they weren't put there by SpamAssassin. If your users are forwarding misclassified messages to a training mailbox as RFC-822 attachments, *you* will have to provide the code that extracts those attachments and passes them to sa-learn.

There is a feature request for building this capability into sa-learn, see https://issues.apache.org/SpamAssassin/show_bug.cgi?id=6928