



Validation results

NetApp Solutions

NetApp
October 20, 2023

Table of Contents

- Validation results 1
 - Pipeline assessment 1
 - Pipeline sentiment metric 1

Validation results

[Previous: Deploying support-center sentiment analysis.](#)

As mentioned in the previous section, errors are propagated throughout the pipeline whenever there are two or more machine learning models running in sequence. For this solution, the sentiment of the sentence is the most important factor in measuring the firm's stock risk level. The speech-to-text model, although essential to the pipeline, serves as the preprocessing unit before the sentiments can be predicted. What really matters is the difference in sentiment between the ground truth sentences and the predicted sentences. This serves as a proxy for the word error rate (WER). The speech-to-text accuracy is important, but the WER is not directly used in the final pipeline metric.

```
PIPELINE_SENTIMENT_METRIC = MEAN(DIFF(GT_sentiment, ASR_sentiment))
```

These sentiment metrics can be calculated for the F1 Score, Recall, and Precision of each sentence. The results can be then aggregated and displayed within a confusion matrix, along with the confidence intervals for each metric.

The benefit of using transfer learning is an increase in model performance for a fraction of data requirements, training time, and cost. The fine-tuned models should also be compared to their baseline versions to ensure the transfer learning enhances the performance instead of impairing it. In other words, the fine-tuned model should perform better on the support center data than the pretrained model.

Pipeline assessment

Test case	Details
Test number	Pipeline sentiment metric
Test prerequisites	Fine-tuned models for speech-to-text and sentiment analysis models
Expected outcome	The sentiment metric of the fine-tuned model performs better than the original pretrained model.

Pipeline sentiment metric

1. Calculate the sentiment metric for the baseline model.
2. Calculate the sentiment metric for the fine-tuned model.
3. Calculate the difference between those metrics.
4. Average the differences across all sentences.

[Next: Videos and demos.](#)

Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.