



Contribution ID: 96

Type: **not specified**

Application of domain-specific language models for quality and technical support in the Food and Beverage Industry

Tuesday, 14 September 2021 17:45 (20 minutes)

Issue Resolution is a critical process in the manufacturing sector to sustain productivity and quality, especially in the Food and Beverage Industry, where aseptic performance is critical. As a leader in this industry, Tetra Pak has built a database regarding quality events reported by Tetra Pak technicians, each containing domain knowledge from experts. In this paper, we present a model framework we have internally developed, which is using a domain-specific language model to address two primary natural language challenges impacting the resolution time:

1. Automatically classify a new reported event to the proper existing class
2. Suggest existing solutions when a new event is being reported, ranked by relevance of the descriptions of the issues (free text documented by the technician)

Our study shows that the language model could benefit from training on domain-specific data compared with those trained on open-domain data. For task 1, the language model is trained on the domain-specific data with an accuracy of over 85%. F1 score average is over 80%. For task 2, the domain-specific deep learning model is combined with a bag-of-words retrieval function-based algorithm to build an advanced search engine with an average precision of 53%.

Keywords

Domain-Specific NLP, Text Classification, Prescriptive Analytics

Special/invited session

Primary authors: Mr LIU, Peng; MONDINO, Chiara; Mr SCHELLENBERG, Noah; Mr BARROSO, Alberto; Ms KYHL, Astrid

Presenters: Mr LIU, Peng; MONDINO, Chiara

Session Classification: Quality 3

Track Classification: Quality