

Preface

Welcome to the Virtual BioCreative VII challenge and workshop, November 08-10, 2021. On behalf of the Organizing Committee, we would like to thank you for your participation and hope you enjoy the workshop.

The BioCreative (Critical Assessment of Information Extraction systems in Biology) challenge evaluation consists of a community-wide effort for evaluating text mining and information extraction systems applied to the biological domain (<https://www.biocreative.org/>). Its aim is to promote the development of text mining and text processing tools which are useful to the communities of researchers, publishers, and database curators in the biological sciences. The main emphasis is on the comparison of methods and the community assessment of scientific progress, rather than on the purely competitive aspects. The first BioCreative was held in 2004, and since then each challenge has ensured the active involvement of the text mining user community in the design of the tracks, preparation of corpus and the testing of interactive systems.

BioCreative VII consisted of five tracks:

Track 1- DrugProt:Text mining drug/chemical-protein interactions: This track explored recognition of chemical-protein entity relations from abstracts. The aim of the DrugProt track is to promote the development and evaluation of systems that are able to automatically detect relations between chemical compounds/drug and genes/proteins.

Track 2- NLM-Chem Track: Full text Chemical Identification and Indexing in PubMed articles: This track aim is to improve automated full-text chemical concept recognition to accelerate manual indexing and curation and advance downstream NLP tasks such as relevant article retrieval. The track consisted of two sub-tasks, focusing on (1) identifying chemicals in full-text articles (i.e. named entity recognition and normalization) and (2) ranking chemical concepts for full-text document indexing.

Track 3- Automatic extraction of medication names in tweets: The goal of this task is to extract the spans that mention a medication or dietary supplement in tweets.

Track 4- COVID-19 text mining tool interactive demo: The goal of this task is to foster the interaction between system developers and potential users to advance in the development of text mining tools that are useful for the research community. This demonstration track focused on tools specifically developed to support COVID-19 research efforts.

Track 5- LitCovid track Multi-label topic classification for COVID-19 literature annotation: The LitCovid track called for a community effort to tackle automated topic annotation for COVID-19 literature as the number of COVID-19-related articles in the literature is growing by about 10,000 articles per month.

A big thanks to all participating teams, panelists, all the chairs, and committee members.

Cecilia Arighi, University of Delaware, USA (on behalf of BioCreative VII organizers)
November 8, 2021

ISBN: 978-0-578-32368-8