Dear Alexander He and Thilanka Munasinghe,

Congratulations!

We are pleased to inform you that your paper BigD630, titled "Chronic Respiratory Disease: Risk Modeling Potential and Limitations" has been accepted as a \*\*short paper\*\* for the 2021 IEEE International Conference on Big Data.

This year we received 486 submissions. After careful review and considering the recommendations from program vice chairs, we accepted 97 regular papers (19.9%), and 96 short papers (19.7%)

We include your paper reviews at the end of this email. Please consider incorporating reviewer comments into your camera-ready version.

Please read the following information about camera ready submission, online registration, video recording & uploading, and student travel awards carefully. If you have any questions, please do not hesitate to contact us.

(1) The final camera-ready paper submission deadline is Nov 15, 2021, pls don’t miss the deadline, otherwise your paper won’t be published in the conference proceedings, Pls follow the URL for the camera-ready paper submission

https://wi-lab.com/cyberchair/2021/bigdata21/index.php

• Regular Papers: 10 pages, including all figures, tables, and references.

• Short Papers: 6 pages, including all figures, tables, and references.

Regular paper: 25 minutes oral presentation (about 20 minutes for talk and 5 minutes for Q and A)

Short paper: 15 minutes (about 11-12 minutes for talk and 3-4 minutes for Q and A)

Extended/poster paper: 15 minutes (about 11-12 minutes for talk and 3-4 minutes for Q and A)

PURCHASING EXTRA PAGES:

The authors could purchase two extra pages if necessary with US$100 per page.

(2) Each paper needs to have at least one full registration in order to get your paper published in the conference proceedings. If you are a student author and you are the only one to register for the paper, then you need to pay the full registration, not the student registration. The author registration deadline is Nov 15, 2021. If you pay by wire-transfer, pls note that wire-transfer normally take about 10 days to actually transfer the fund from your bank account to the IEEE BIBM 2021 account, make sure your payment is received by Nov 15, otherwise your paper won’t be published in the conference proceedings. Pls follow the registration link from

https://bigdataieee.org/BigData2021/ for the online registration

(3) IEEE BigData 2021 is taking place Virtually from Dec 15-18, 2021. Pls read the video recording instruction https://docs.google.com/presentation/d/17urdXFUcF6mFnepabBPj07Cd9Dk9MjCL0qW6waBb46E/edit?usp=sharing and then submit the video through this link https://ieeebigdata2021videosubmission.paperform.co/

Please note the vide submission deadline is Nov 24, 2021.

(4) The conference schedule will be announced around late November. WE will email you the login in instruction to the virtual platform to attend the conference a few days before the event. So everyone wants to attend the conference must register the conference in order to login in the virtual conference platform to attend the conference.

(5) Thanks for the student travel grant from National Science Foundation of USA, the conference will offer about 75 student travel awards. Per NSF policy, only the students from institutes in USA are eligible to apply. If you are a student author from a USA institute, please email Dr. Min Shi at mins@wustl.edu the following materials before the deadline Nov 21, 2021: your name, affiliation, supporting letter from your advisor or dept chair, paper id and title, registration receipt

Sincerely yours,

Yixin Chen, Washington University at St Louis, USA

Heiko Ludwig, IBM Almaden Research Center, USA

Yicheng Tu, South Florida University, USA

============================================================================

--======== Review Reports ========--

The review report from reviewer #1:

\*1: Is the paper relevant to Bigdata?

[\_] No

[X] Yes

\*2: How innovative is the paper?

[\_] 5 (Very innovative)

[\_] 4 (Innovative)

[X] 3 (Marginally)

[\_] 2 (Not very much)

[\_] 1 (Not)

[\_] 0 (Not at all)

\*3: How would you rate the technical quality of the paper?

[\_] 5 (Very high)

[\_] 4 (High)

[X] 3 (Good)

[\_] 2 (Needs improvement)

[\_] 1 (Low)

[\_] 0 (Very low)

\*4: How is the presentation?

[\_] 5 (Excellent)

[X] 4 (Good)

[\_] 3 (Above average)

[\_] 2 (Below average)

[\_] 1 (Fair)

[\_] 0 (Poor)

\*5: Is the paper of interest to Bigdata users and practitioners?

[\_] 3 (Yes)

[X] 2 (May be)

[\_] 1 (No)

[\_] 0 (Not applicable)

\*6: What is your confidence in your review of this paper?

[X] 2 (High)

[\_] 1 (Medium)

[\_] 0 (Low)

\*7: Overall recommendation

[\_] 5 (Strong Accept: top quality)

[\_] 4 (Accept: a regular paper)

[X] 3 (Weak Accept: could be a poster or a short paper)

[\_] 2 (Weak Reject: don't like it, but won't argue to reject it)

[\_] 1 (Reject: will argue to reject it)

[\_] 0 (Strong Reject: hopeless)

\*8: Detailed comments for the authors

Te paper makes a detailed study on factors that cause chronic respiratory disease, using data from different applying standard machine learning methods for feature analysis.

The domain specific findings (what factors actually correlate with CRD) are probably of little interest to the Big Data community, but rather to domain specialists in environmental research and public health. However, the paper gives a nice illustration of how ML can be concretely applied in a practical setting and what are limitations of ML applied to real-world data.

========================================================

The review report from reviewer #2:

\*1: Is the paper relevant to Bigdata?

[\_] No

[X] Yes

\*2: How innovative is the paper?

[\_] 5 (Very innovative)

[X] 4 (Innovative)

[\_] 3 (Marginally)

[\_] 2 (Not very much)

[\_] 1 (Not)

[\_] 0 (Not at all)

\*3: How would you rate the technical quality of the paper?

[\_] 5 (Very high)

[\_] 4 (High)

[\_] 3 (Good)

[X] 2 (Needs improvement)

[\_] 1 (Low)

[\_] 0 (Very low)

\*4: How is the presentation?

[\_] 5 (Excellent)

[\_] 4 (Good)

[X] 3 (Above average)

[\_] 2 (Below average)

[\_] 1 (Fair)

[\_] 0 (Poor)

\*5: Is the paper of interest to Bigdata users and practitioners?

[\_] 3 (Yes)

[X] 2 (May be)

[\_] 1 (No)

[\_] 0 (Not applicable)

\*6: What is your confidence in your review of this paper?

[\_] 2 (High)

[X] 1 (Medium)

[\_] 0 (Low)

\*7: Overall recommendation

[\_] 5 (Strong Accept: top quality)

[\_] 4 (Accept: a regular paper)

[X] 3 (Weak Accept: could be a poster or a short paper)

[\_] 2 (Weak Reject: don't like it, but won't argue to reject it)

[\_] 1 (Reject: will argue to reject it)

[\_] 0 (Strong Reject: hopeless)

\*8: Detailed comments for the authors

The paper analyzes risks of Chronic Respiratory Diseases (CRD) including emphysema, asthma, bronchiectasis, and Chronic Obstructive Pulmonary Disease. It takes into consideration ambient air pollution and changes in temperature and humidity and analyzes how well machine learning can be used for the prediction of the CRDs. This required integration of information about mortality, fine particulate matter, population, median income, climate variable, fires and air quality. There were many limitations and approximations because of limited data availability. The paper discusses their feature engineering in a lot of detail, however at this point it lacks clear conclusions or recommendations. Overall, it’s a good project but at relatively early stage.

========================================================