COVID Information Commons (CIC) Research Lightning Talk

Transcript of a Presentation by Ioannis Paschalidis (Boston University), April 24, 2023

Title: Predictive Models of COVID-19 Severity and Patient Outcomes

Ioannis Paschalidis CIC Database Profile

NSF Award #: 2200052

YouTube Recording with Slides

Spring 2023 CIC Webinar Information

Transcript Editor: Shikhar Johri

Transcript

स्लाइड *1*

आयोनिस पास्चालिडिस:

बहुत बहुत धन्यवाद और वक्ताओं के एक महान सेट के साथ इस रोमांचक दोपहर का हिस्सा बनने के लिए मुझे आमंत्रित करने के लिए धन्यवाद। मैं COVID-19 गंभीरता और रोगी परिणामों के भविष्य कहनेवाला मॉडल के बारे में बात करूंगा जो हमने बोस्टन विश्वविद्यालय में विकसित किए हैं।

स्लाइड *2*

बस थोड़ा सा संदर्भ देने के लिए - हम महामारी से पहले कुछ समय से विभिन्न मॉडलों पर काम कर रहे हैं जो बीमारी और महत्वपूर्ण प्रमुख घटनाओं की भविष्यवाणी करते हैं, उदाहरण के लिए, अस्पताल में भर्ती, और ऐसे मॉडल भी जो उपचार निर्धारित करते हैं। इसलिए जब महामारी शुरू हुई, तो हम बाकी वैज्ञानिक समुदाय की तरह जुटे, और हमने सहयोगियों का एक अपेक्षाकृत बड़ा नेटवर्क भी जुटाया, इसलिए हमने विभिन्न डेटा सेटों तक पहुंच प्राप्त की।

आप इस स्लाइड में हमारे पास मौजूद विभिन्न डेटा सेट देख सकते हैं। डेटा सेट जो मैसाचुसेट्स से स्थानीय थे - मैसाचुसेट्स में दो अलग-अलग अस्पताल नेटवर्क से। एक मास जनरल ब्रिघम नेटवर्क से था, पांच अलग-अलग अस्पतालों से लगभग 2,500 मामले और फिर बोस्टन मेडिकल सेंटर से लगभग 7,000 मामलों का एक और अपेक्षाकृत बड़ा समूह। हमें वुहान, चीन से भी कुछ मामले मिले जो स्पष्ट रूप से महामारी का उद्गम स्थल था। और अंत में हम कुछ बड़े राष्ट्रीय डेटासेट तक पहुंच प्राप्त करने में सक्षम थे - इसलिए ब्राजील से एक डेटा सेट और मेक्सिको से एक और डेटा सेट।

इस बहुत ही छोटी प्रस्तुति में मैं सबसे हाल के काम पर अधिक ध्यान केंद्रित करूंगा जिसने डेटासेट पर विचार किया - बोस्टन मेडिकल सेंटर के रोगियों की श्रृंखला। बोस्टन मेडिकल सेंटर बीयू मेडिकल स्कूल से

संबद्ध शिक्षण अस्पताल है और एक सेफ्टीनेट अस्पताल भी है। जैसा कि आप देखेंगे, निष्कर्षों में इसके कुछ दिलचस्प निहितार्थ हैं जो हम प्राप्त करने में सक्षम थे।

स्लाइड *3*

इसलिए हमें लगभग 7,000 रोगियों के पूरे 2020 बीएमसी समूह तक पहुंच मिली। वे मरीज थे जिन्होंने COVID-19 के लिए सकारात्मक परीक्षण किया था। बस कुछ मोटे आंकड़े - लगभग 20 प्रतिशत भर्ती कराए गए थे। भर्ती किए गए लोगों में से लगभग 23 या तो आईसीयू में भर्ती थे। आईसीयू में भर्ती लोगों में से लगभग 60 या 58.7 प्रतिशत इंटुबैट किए गए थे। इंटुबैट किए गए लोगों से, लगभग 70 प्रतिशत दुर्भाग्य से इसे नहीं बना पाए। हमारे पास इन रोगियों के बारे में बहुत सारी जानकारी थी, जिसमें जनसांख्यिकी, अस्पताल में रहने के दौरान उनके महत्वपूर्ण कार्य, रेडियोलॉजी रिपोर्ट, उनका चिकित्सा इतिहास, कोई भी लक्षण, कोई प्रयोगशाला परिणाम, कोई दवा, और यहां तक कि अवसाद की स्थिति, ज़िप कोड के बारे में भी जानकारी थी, और हमारे पास उस समय अस्पताल के अधिभोग के बारे में जानकारी थी कि इनमें से प्रत्येक रोगी को देखा गया था। हमारे पास स्वास्थ्य के सामाजिक निर्धारकों के बारे में भी जानकारी थी। बीएमसी थ्राइव प्रोग्राम नामक एक कार्यक्रम चलाता है कि अस्पताल के साथ मुठभेड़ करने वाले हर किसी को एक सर्वेक्षण दिया जाता है, जहां हमने उनसे आवास, भोजन, परिवहन, देखभाल करने में मदद, दवाओं, दवाओं तक पहुंच प्राप्त करने और दवाओं, शिक्षा और रोजगार के लिए भुगतान करने में मदद सहित विभिन्न क्षेत्रों में जरूरतों के बारे में पूछा। बहुत सारी जानकारी सारणीबद्ध प्रारूप में थी जिसे एआई मशीन सीखने के तरीकों द्वारा अधिक आसानी से नियंत्रित किया जाता है, लेकिन अस्पताल में रहने के दौरान विशेष रूप से रेडियोलॉजी रिपोर्ट और अन्य रिपोर्टों में काफी जानकारी थी, जो सिर्फ कथाएं थीं - चिकित्सकों द्वारा रिपोर्ट। पाठ से उचित निष्कर्ष निकालने के लिए हमें काफी प्राकृतिक भाषा प्रसंस्करण का उपयोग करना पड़ा। हमने जो किया वह हमने मजबूत व्याख्यात्मक मॉडल का एक सेट विकसित किया है जो अस्पताल में भर्ती, आईसीयू प्रवेश, यांत्रिक वेंटिलेशन और मृत्यु की भविष्यवाणी कर सकता है। मैं आपको कुछ उदाहरण दिखाता हूँ। मैं समय के हित में फिर से पूरी तरह से नहीं बोलूंगा।

स्लाइड *4*

सबसे पहले, हमने जो किया वह प्रत्येक रोगी के लिए है, हमने उस समय से एक समयरेखा का निर्माण किया जब हम रोगी के सकारात्मक परीक्षण के बारे में जानते थे और इस रोगी के बारे में ब्याज की घटना के बिंदु तक जानकारी रखते थे, चाहे वह हो, मान लीजिए, आईसीयू प्रवेश या यांत्रिक वेंटिलेशन। यहां ब्याज का परिणाम है, और फिर समय पर वापस देखकर, हमने इन डेटा बकेट को बनाया। हमने ब्याज की घटना के समय से ठीक पहले उपलब्ध किसी भी जानकारी को छोड़ दिया क्योंकि हम ऐसे मॉडल चाहते थे जो भविष्यवाणी कर सकें कि भविष्य में क्या होगा। शायद एक प्रथम वर्ष का मेडिकल छात्र भी आईसीयू में जाने वाले रोगी की पहचान कर सकता है, इसलिए हम चाहते थे कि मॉडल पहले की जानकारी के साथ उस भविष्यवाणी को करें। आप देखेंगे कि हमने उन मॉडलों के विभिन्न संस्करण प्राप्त किए हैं जिनमें भविष्यवाणी करने के लिए मॉडल द्वारा उपयोग की जाने वाली जानकारी के संदर्भ में अलग-अलग कटऑफ थे। एक और कारण है कि हमने इन समय बाल्टी को बनाया है, हम अस्पताल में रोगी की प्रगति के गतिशील विकास को पकड़ना चाहते थे, उदाहरण के लिए, विटल्स का गतिशील विकास। हम समझ गए कि यह काफी महत्वपूर्ण था। इसलिए केवल विटल्स के स्नैपशॉट को देखने के बजाय, मान लें कि कुछ विशिष्ट समय पर और आगे की भविष्यवाणी करने के लिए उस जानकारी का उपयोग करते हुए, सटीक मूल्य महत्वपूर्ण हैं लेकिन रुझान भी महत्वपूर्ण हैं। चिकित्सक, जब वे रोगियों को देखते हैं, तो वे रोगी में महत्वपूर्ण बातों के रुझानों को देखते हैं। इसलिए हमने एक मॉडल विकसित किया जिसमें एलएसटीएम प्रकार के नेटवर्क और एक ट्रांसफार्मर आर्किटेक्चर सहित कुछ काफी परिष्कृत गहन शिक्षण पद्धतियों का उपयोग किया गया, जो समय के विभिन्न बिंदुओं पर छह महत्वपूर्ण संकेतों को इनपुट के रूप में लेते थे - और एक स्कोर का उत्पादन करते थे। उस स्कोर ने विटल्स के गतिशील विकास पर कब्जा कर लिया और

उस महत्वपूर्ण स्कोर का उपयोग तब एक पहनावा मॉडल में किया गया था जो ब्याज के परिणाम के लिए भविष्यवाणी करने का प्रयास कर रहा था।

स्लाइड *5*

उदाहरण के लिए, अस्पताल में भर्ती होने की भविष्यवाणियां। आप देख सकते हैं कि ये काफी सटीक हैं। कुछ बेहतरीन मॉडल आपको 92 प्रतिशत देते हैं - यह वक्र के नीचे के क्षेत्र में है। आप इसे मॉडल की सटीकता के उपाय के रूप में सोच सकते हैं। सबसे अच्छा 100 प्रतिशत है, एक यादृच्छिक अनुमान आपको 50 प्रतिशत देगा। तो 92 प्रतिशत काफी अच्छा प्रदर्शन है। आप यहां देख सकते हैं कि हमारे द्वारा विकसित कुछ रैखिक मॉडलों से हमें कुछ ऐसे कारक भी मिले जो अस्पताल में भर्ती होने की भविष्यवाणी करने में महत्वपूर्ण थे। नीले रंग में, आप कुछ चर देखते हैं जो कुछ पहले की स्वास्थ्य स्थितियों से जुड़े होते हैं, उदाहरण के लिए, अस्पताल में भर्ती होने के साथ अत्यधिक सहसंबद्ध होते हैं। आप यह भी देखेंगे कि अस्पताल की अधिभोग, यदि यह अधिक थी, तो रोगी के अस्पताल में भर्ती होने की संभावना कम हो गई। इसके अलावा, आप स्वास्थ्य के दो सामाजिक निर्धारक देखेंगे: भोजन की आवश्यकता और परिवहन की आवश्यकता। ये दोनों अस्पताल में भर्ती होने के फैसले में योगदान दे रहे थे। उन जरूरतों वाले मरीजों को अस्पताल में भर्ती होने की अधिक संभावना थी। मैं स्वास्थ्य के इन सामाजिक निर्धारकों की भूमिका पर जोर देना चाहूंगा। यह कुछ ऐसा है जिसे हमने अन्य डेटासेट में भी देखा, विशेष रूप से ब्राज़ीलियाई डेटासेट में, जो एक राष्ट्रीय डेटा सेट था और हमने पाया कि सामाजिक-जनसांख्यिकीय कारक अस्पताल में भर्ती होने के निर्णयों को प्रभावित कर रहे थे।

हमने यह भी पाया कि मॉडल - भोला मॉडल जो उत्पादन करने में सक्षम है - वास्तव में पक्षपाती है। आप यहां देख सकते हैं कि मॉडल काले व्यक्तियों और सफेद व्यक्तियों के लिए नमूनों से बाहर कैसे प्रदर्शन करता है। काले व्यक्तियों के लिए मॉडल की झूठी सकारात्मक दर मॉडल के रूप में दोगुनी थी - सफेद व्यक्तियों के लिए झूठी सकारात्मक दर - भले ही हम दौड़ के लिए नियंत्रण कर रहे थे, हम सामाजिक-जनसांख्यिकीय कारकों के लिए नियंत्रण कर रहे थे, हम स्वास्थ्य के सामाजिक निर्धारकों के लिए नियंत्रण कर रहे थे। इसके बावजूद, मॉडल यह भविष्यवाणी करने के लिए बहुत अधिक उत्सुक था कि एक श्वेत व्यक्ति की तुलना में अश्वेत व्यक्ति को अस्पताल में भर्ती कराया जाएगा। इसी तरह, यह एक काले व्यक्ति की तुलना में एक सफेद व्यक्ति के लिए एक झूठी नकारात्मक भविष्यवाणी करने की सबसे अधिक संभावना थी, जो बताता है कि वे डेटा में स्पष्ट रूप से छिपी हुई विशेषताएं हैं जो हमें दिखाई नहीं दे रही हैं, शायद संरचनात्मक पूर्वाग्रह और अन्य कारकों को दर्शाती हैं जो मॉडल को पक्षपातपूर्ण भविष्यवाणी करते हैं। ऐसे तरीके हैं और हमने उन्हें एक पेपर में संबोधित किया है जिसे हमने प्रकाशित किया है कि कोई इन कारकों के लिए कैसे सही हो सकता है और ऐसे मॉडल तैयार कर सकता है जिनमें इस तरह का पूर्वाग्रह नहीं है।

स्लाइड *6*

ये आईसीयू भविष्यवाणी पर कुछ परिणाम हैं - आईसीयू प्रवेश की भविष्यवाणी करना - अस्पताल में प्रवेश और आईसीयू में प्रवेश के बीच औसत अंतर, कम से कम हमारे डेटा सेट में, लगभग चार घंटे था। हम अलग-अलग कटऑफ का उपयोग करेंगे। यदि आप नवीनतम जानकारी का उपयोग करते हैं, तो आपको 93-95 प्रतिशत के आदेश पर एयूसी के साथ काफी सटीक मॉडल मिलते हैं। यदि आप उस जानकारी को काटना शुरू करते हैं जिसका आप उपयोग करने जा रहे थे, तो 12 घंटे पहले, मॉडल का प्रदर्शन लगभग 86 प्रतिशत तक गिर जाता है। 24 घंटे पहले, मॉडल का प्रदर्शन लगभग 80 प्रतिशत तक गिर जाता है। मैं जिस बात पर जोर देना चाहता हूं वह यह है कि हमने इन मॉडलों की तुलना कुछ मानक मॉडलों से की है जो आईसीयू प्रवेश की भविष्यवाणी करते हैं, कुछ प्रसिद्ध सेप्सिस मॉडल हैं, समाचार, क्यूएसओएफए, उन्हें कहा जाता है, और ये इस मामले में काफी गलत हैं, यह दर्शाता है कि आईसीयू भविष्यवाणी के लिए मानक

मॉडल, कम से कम COVID मामलों में, आईसीयू प्रवेश की भविष्यवाणी करने में विफल रहते हैं। यह बीमारी के एक अद्वितीय हस्ताक्षर को इंगित करता है। यहां, आपको कुछ चर मिलते हैं जो फिर से आईसीयू प्रवेश के साथ परिणाम के साथ अत्यधिक सहसंबद्ध थे। हमारे लिए दिलचस्प बात यह थी कि यह महत्वपूर्ण स्कोर जो हमने उत्पादित किया था जिसने विटल्स के गतिशील विकास को पकड़ लिया था, पूरी कहानी को काफी हद तक बताता है। कुछ अन्य चर या कुछ प्रयोगशाला चर (एलडीएच, सीआरपी) हैं जिन्हें अन्य अध्ययनों द्वारा पहचाना गया है जो भी योगदान दे रहे हैं, लेकिन अगर कोई सिर्फ महत्वपूर्ण के गतिशील विकास को लेता है, तो यह आईसीयू प्रवेश कहानी को बहुत अधिक बताता है।

स्लाइड *7*

अंत में हमने कई कैलकुलेटर तैयार किए जिन्हें हमने वेब पर उपलब्ध कराया। यह, मैं समझता हूं, महामारी के शुरुआती चरणों में मास जनरल अस्पताल में हमारे सहयोगियों और सहयोगियों द्वारा उपयोग किया गया था। कुछ प्रमुख चरों को इनपुट करना और भविष्यवाणी प्राप्त करना बहुत आसान था, उदाहरण के लिए, आईसीयू प्रवेश के लिए, या किसी रोगी की यांत्रिक इंटुबैषेण आवश्यकता के लिए, और हमने पाया कि, आप जानते हैं, हमें मामलों को खोजने के लिए चुनौती दी गई थी - मॉडल क्या है मुझे बता रहा है कि बहुत अनुभवी चिकित्सक केवल रोगी को देखकर संभावित रूप से भविष्यवाणी नहीं कर सकता है। यहां कुछ मामले हैं, और कई अन्य हैं, जहां रोगियों को भर्ती कराया गया था, वे कुछ दिनों के लिए स्थिर थे, उनके नैदानिक दृष्टिकोण में कुछ भी नहीं सुझाया गया था कि इस रोगी की स्थिति बिगड़ने वाली थी, लेकिन तब मॉडल सक्षम था - प्रवेश पर - कुछ विशेष प्रयोगशाला परिणामों के आधार पर, भविष्यवाणी करें कि रोगी को आईसीयू देखभाल की आवश्यकता होगी। ये दो मामले हैं जिनमें उन मामलों का ब्यौरा है।

स्लाइड *8*

तो यह मुझे एक निष्कर्ष पर लाता है। बेशक, मैं सिर्फ प्रस्तुत कर रहा हूं। ऐसे कई लोग हैं जिन्होंने इस काम में योगदान दिया है और मैं उन्हें धन्यवाद देना चाहता हूं, जिसमें मेरे समूह के छात्र भी शामिल हैं, लेकिन मास जनरल ब्रिघम बोस्टन मेडिकल सेंटर और कुछ अन्य क्षेत्रों में हमारे सहयोगी भी हैं जहां से हम डेटा प्राप्त करने में सक्षम हैं। आपका ध्यान देने और सत्र के अंत में प्रश्नों की प्रतीक्षा करने के लिए बहुत-बहुत धन्यवाद।