**Memprediksi Kemungkinan Infeksi COVID-19**

**menggunakan Sistem Logika Fuzzy**

**Pengumpulan dan Prapemrosesan Data**

Karena proyek ini didasarkan pada situasi yang saat ini berkembang dan berkembang, luasnya data tersedia. Sebagian besar data difokuskan pada epidemiologi, bagaimanapun, dan detail yang relatif sedikit adalah tersedia pada gejala yang tepat yang dialami oleh pasien itu sendiri.

Tujuan utama selama pengumpulan data adalah untuk menemukan data yang menghubungkan timbulnya gejala dan tingkat keparahan gejala-gejala tersebut dengan kemungkinan seseorang tertular COVID19 yang sangat menular. Namun, upaya ini terhambat oleh kurangnya informasi tentang tingkat keparahan. Informasi seperti tingkat demam yang tepat atau laju pernapasan dijaga kerahasiaannya oleh rumah sakit, karena dirahasiakan informasi medis milik pasien mereka.

Namun, ada banyak makalah penelitian medis yang diterbitkan selama beberapa bulan terakhir, yang disusun ke dalam Covid-19 Open Research Dataset Challenge [7]. Makalah-makalah ini berisi data tentang penyebaran epidemi, gejala yang dilaporkan oleh pasien, tingkat keparahan gejala-gejala itu, faktor-faktor yang berkontribusi terhadap kasus yang parah [8], dan kemanjuran berbagai perawatan dan tindakan medis. Makalah-makalah ini sangat penting untuk membangun Basis Aturan FLS kami.

Untuk membangun fungsi keanggotaan kami, kami malah melihat himpunan data yang berisi data tentang yang diketahui infeksi dan gejala umum yang dilaporkan. Kami menemukan dua kumpulan data tersebut [15, 16], yang berisi informasi seperti tanggal masuk, lokasi geografis, usia dan gejala yang dilaporkan. Oleh prapemrosesan dan pembuatan bagan data, kami dapat mencapai kesimpulan mengenai yang paling umum gejala yang dilaporkan. Setelah menggabungkan dua himpunan data dan menghapus tupel duplikat dan tidak perlu tupel (gejala yang dilaporkan oleh mereka yang dites negatif COVID-19), kami memiliki ukuran sampel 1514 Pasien. 1514 pasien ini semuanya dilaporkan positif COVID-19. Sedangkan data mengenai mereka yang menunjukkan gejala tetapi dites negatif untuk COVID-19 akan sangat berguna, data seperti itu tidak tersedia.



Demam sebagai Gejala

Batuk sebagai Gejala





Otot, Dada dan Nyeri lainnya

Kesulitan Bernapas

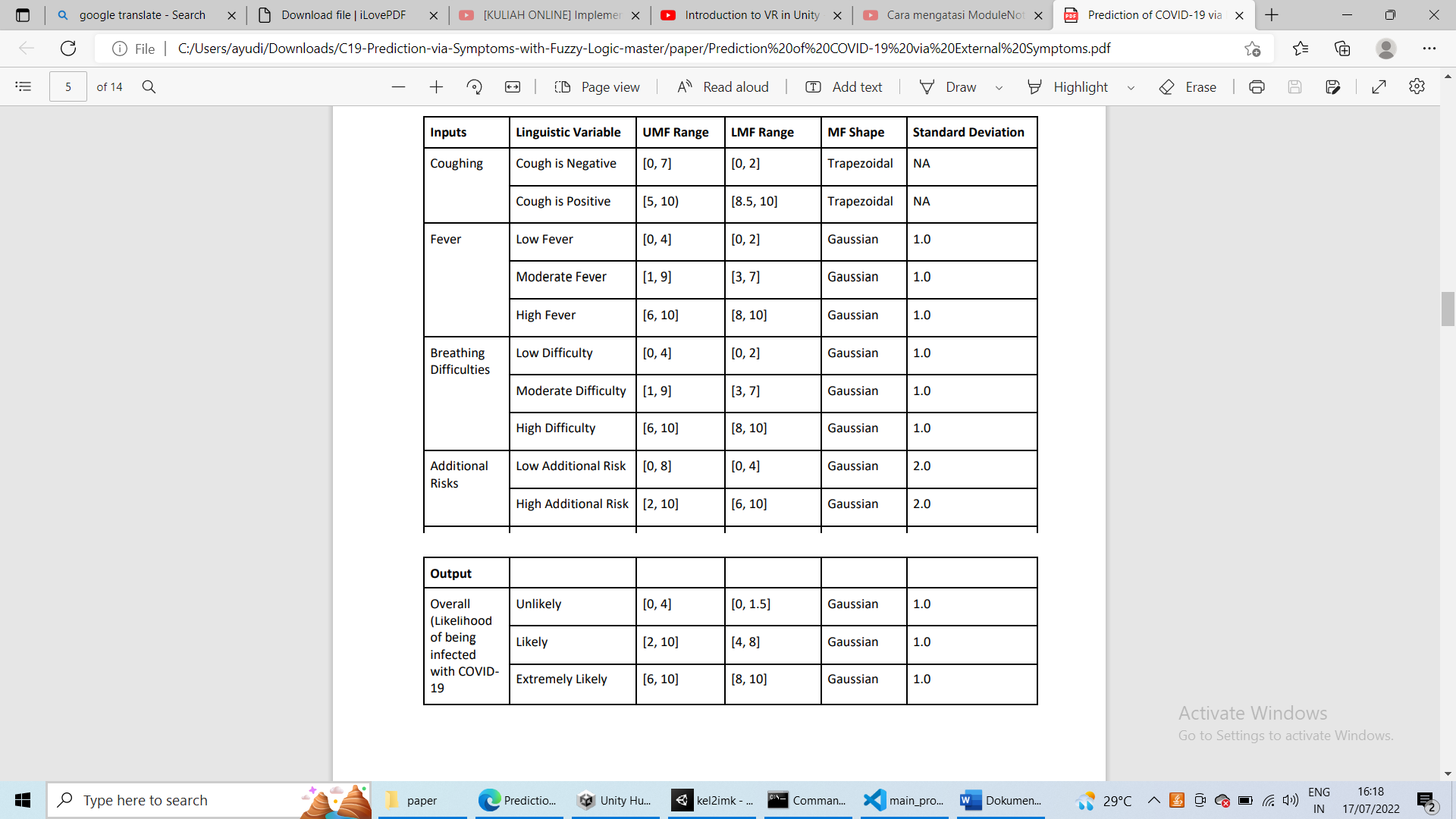
**Desain Sistem Logika Fuzzy**

Ada tiga tahap utama dalam desain FLS kami: Fuzzification, atau menentukan input dan fungsi keanggotaan untuk masing-masing anteseden; Pembuatan Basis Aturan; Defuzzifikasi, atau mengevaluasi output dari sistem.

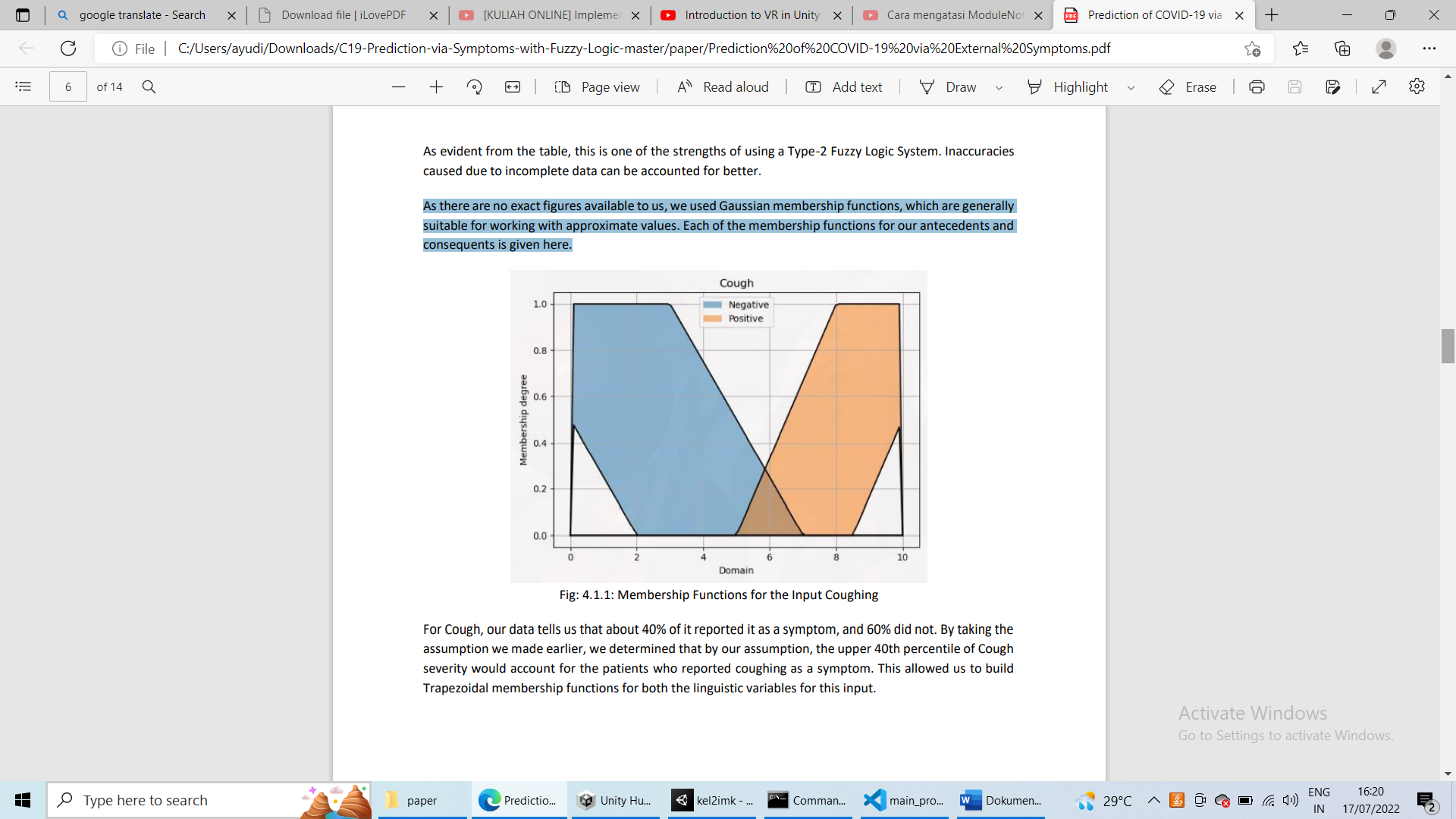
1. Menentukan Fungsi Keanggotaan

Berbagai macam gejala dilaporkan oleh individu yang menderita COVID-19. Kami menggabungkan semua gejala dan faktor risiko yang berkontribusi ke dalam empat kategori umum: Batuk, Demam, Kesulitan Bernapas dan Risiko Tambahan. Karena kurangnya informasi konkret tentang tingkat keparahan setiap gejala, kami terpaksa menggunakan bundaran pendekatan membuat asumsi utama bahwa pasien melaporkan gejala tertentu hanya jika itu cukup parah. Secara umum, fungsi keanggotaan kami dibangun dengan mengekstrapolasi informasi dari makalah dan himpunan data lainnya.

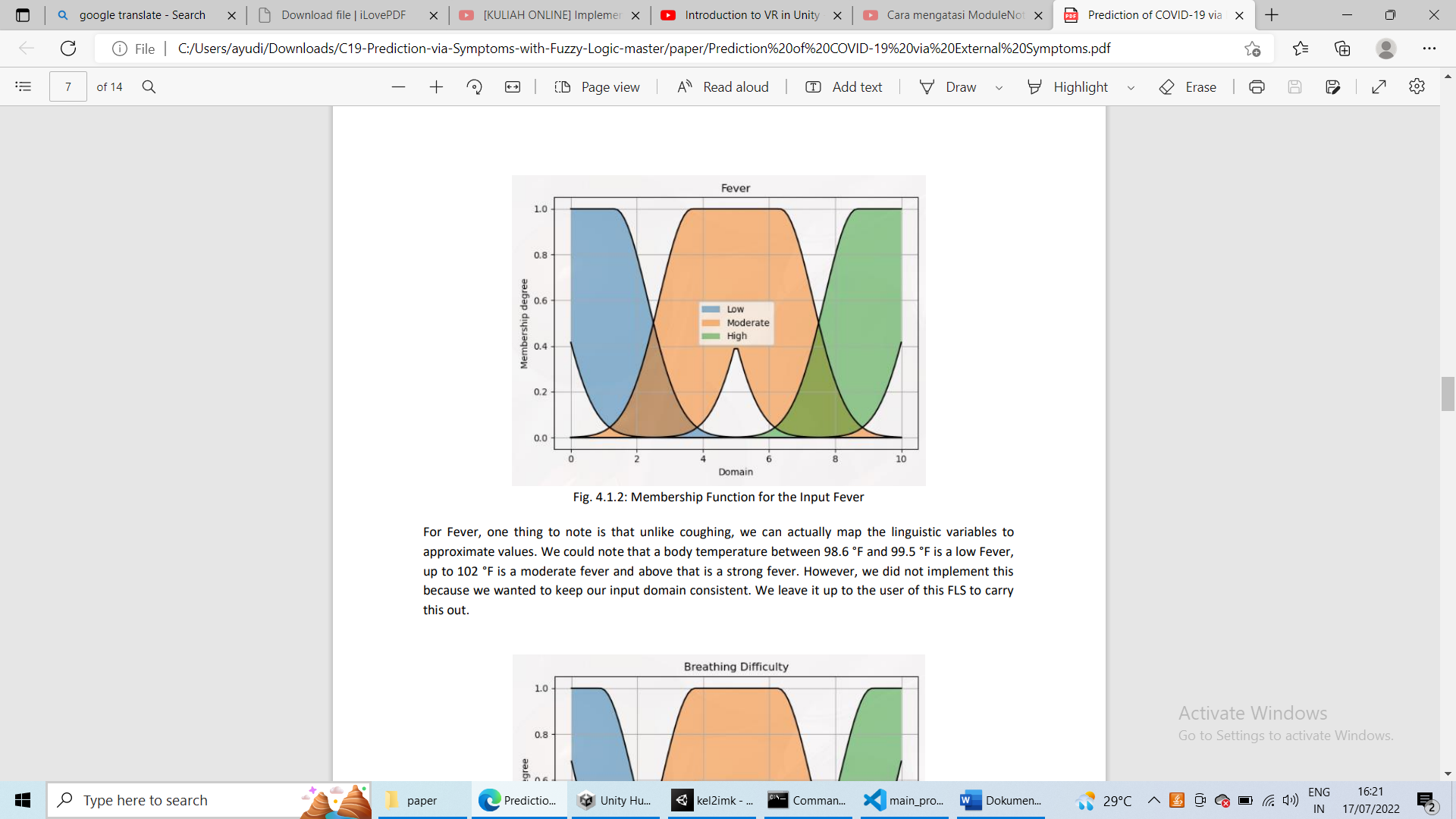
Semua fungsi keanggotaan berada dalam domain [0, 10]. Domain ini mewakili input dari a skala geser yang dapat dimasuki pengguna. Tabel yang diberikan di bawah ini menunjukkan perkiraan domain dari masing-masing domain kami variabel linguistik. Perhatikan bahwa domain ini dipisahkan ke dalam Keanggotaan Atas dan Bawah Fungsi. Selain itu, simpangan baku fungsi gaussian diberikan di sini. Dengan ini informasi informasi matematika dari kurva aktual dapat ditentukan.



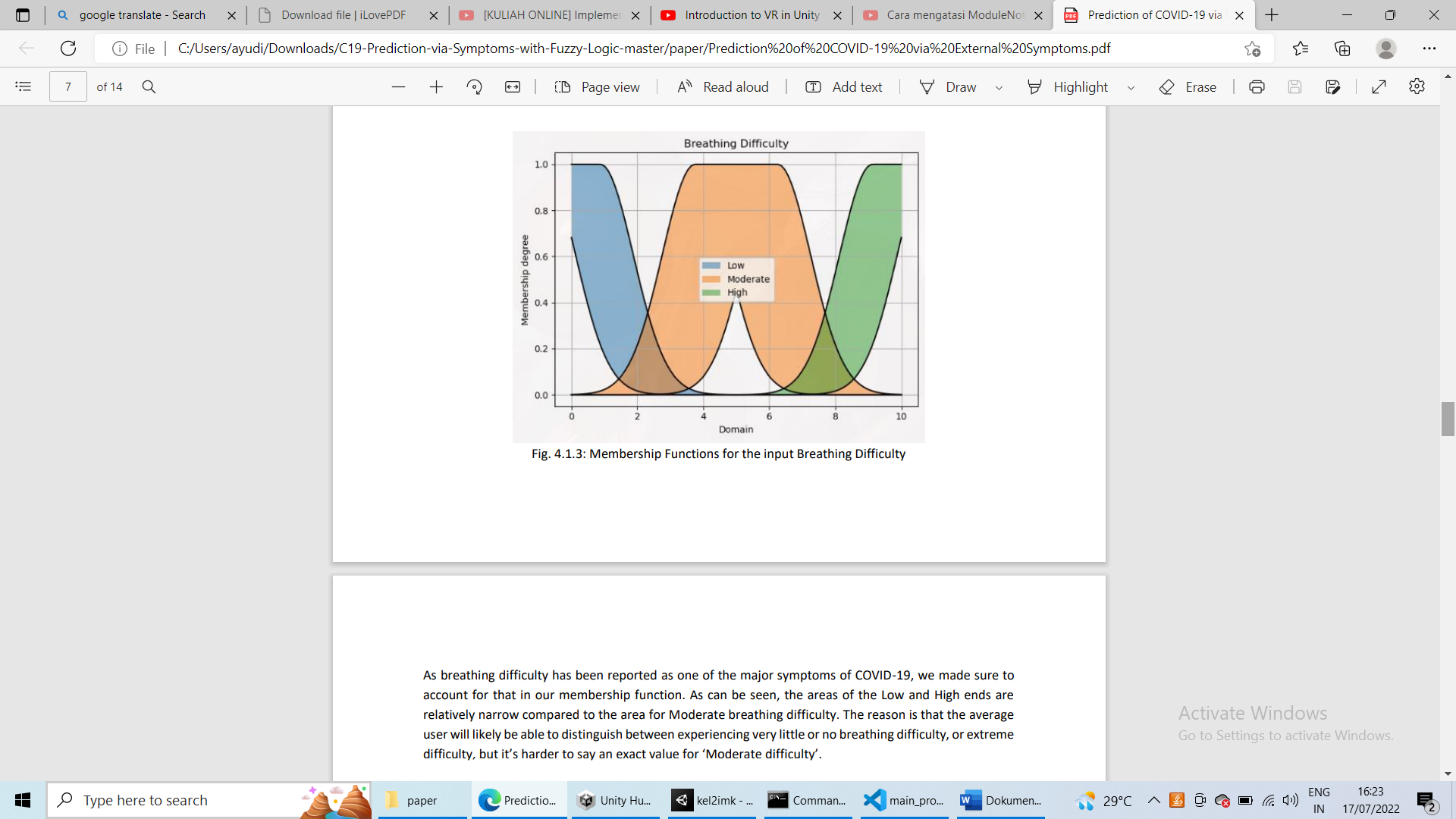
Terbukti dari tabel, ini adalah salah satu kekuatan menggunakan Sistem Logika Fuzzy Tipe-2. Ketidakakuratan disebabkan karena data yang tidak lengkap dapat dipertanggungjawabkan dengan lebih baik. Karena tidak ada angka pasti yang tersedia bagi kami, kami menggunakan fungsi keanggotaan Gaussian, yang umumnya cocok untuk bekerja dengan nilai perkiraan. Masing-masing fungsi keanggotaan untuk anteseden kami dan konsekuensi diberikan di sini.



Untuk Cough, data kami memberi tahu kami bahwa sekitar 40% darinya melaporkannya sebagai gejala, dan 60% tidak. Dengan mengambilasumsi yang kami buat sebelumnya, kami menentukan bahwa dengan asumsi kami, persentil ke-40 atas Batuk tingkat keparahan akan memperhitungkan pasien yang melaporkan batuk sebagai gejala. Ini memungkinkan kami untuk membangun Keanggotaan trapesium berfungsi untuk kedua variabel linguistik untuk input ini.

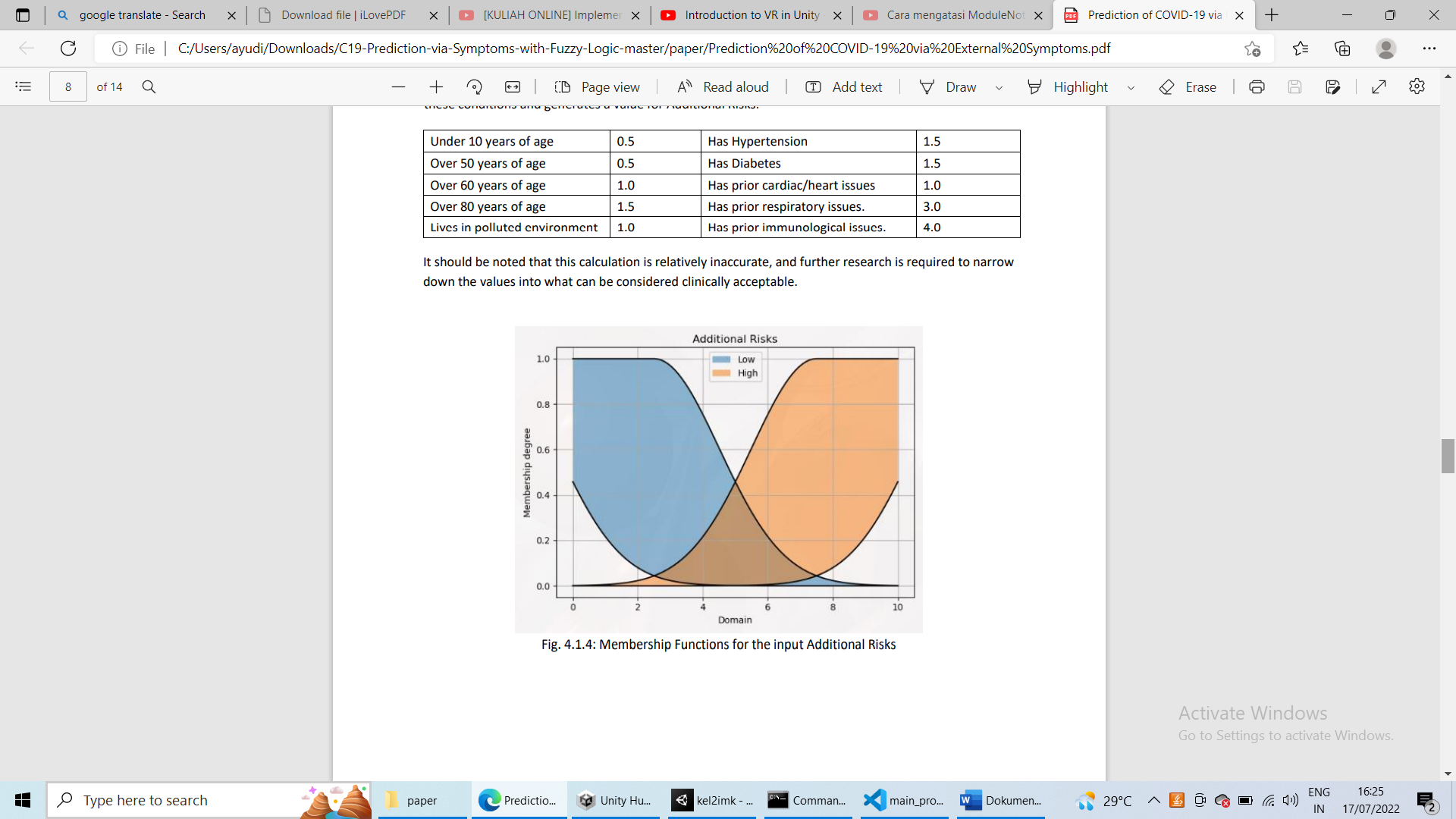


Untuk Fever, satu hal yang perlu diperhatikan adalah bahwa tidak seperti batuk, kita sebenarnya dapat memetakan variabel linguistik ke nilai perkiraan. Kita dapat mencatat bahwa suhu tubuh antara 98,6 °F dan 99,5 °F adalah Demam rendah, hingga 102 °F adalah demam sedang dan di atasnya adalah demam yang kuat. Namun, kami tidak menerapkan ini karena kami ingin menjaga domain input kami tetap konsisten. Kami menyerahkannya kepada pengguna FLS ini untuk dibawa ini keluar.

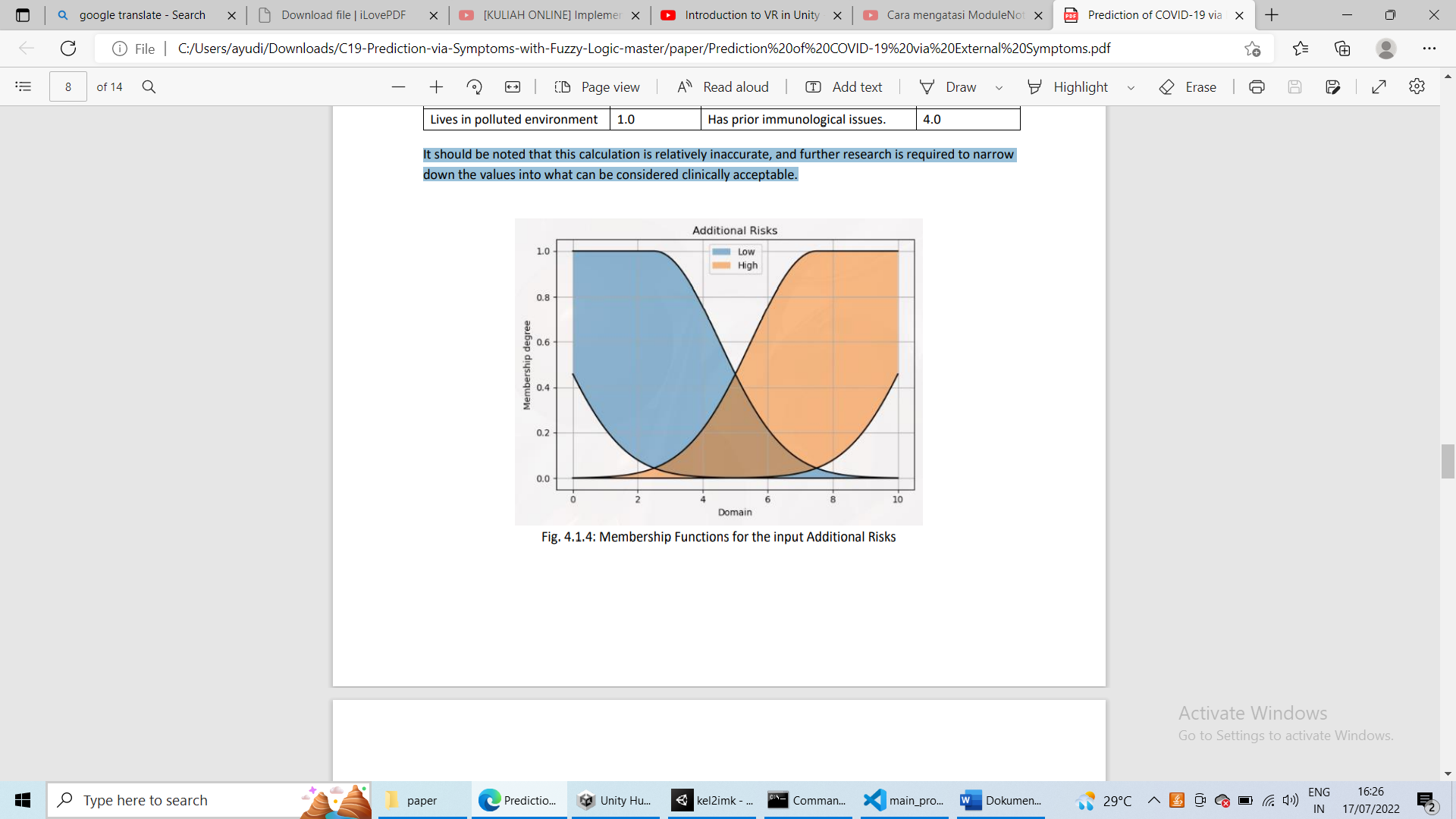


Karena kesulitan bernapas telah dilaporkan sebagai salah satu gejala utama COVID-19, kami memastikan untuk memperhitungkannya dalam fungsi keanggotaan kami. Seperti dapat dilihat, area ujung Rendah dan Tinggi adalah relatif sempit dibandingkan dengan area untuk kesulitan bernapas sedang. Alasannya adalah rata-rata pengguna kemungkinan akan dapat membedakan antara mengalami sangat sedikit atau tidak ada kesulitan bernapas, atau ekstrem kesulitan, tetapi lebih sulit untuk mengatakan nilai yang tepat untuk 'Kesulitan sedang'.

Akhirnya, Risiko Tambahan adalah salah satu kesulitan utama. Risiko dan bahaya infeksi COVID-19 adalah diperbesar oleh berbagai komorbiditas seperti masalah kardiovaskular atau masalah pernapasan. Kami memilih paling umum dari komorbiditas ini: Hipertensi [9], Diabetes, lainnya Masalah Kardiovaskular, ada Masalah Pernapasan, dan Masalah Imunologis. Dengan mempertimbangkan Usia dan Lingkungan juga, yang telah dieksplorasi dalam beberapa makalah medis [10], kami membangun fungsi sekunder yang mengambil input dari kondisi ini dan menghasilkan nilai untuk Risiko Tambahan.



Perlu dicatat bahwa perhitungan ini relatif tidak akurat, dan penelitian lebih lanjut diperlukan untuk mempersempit turun nilai ke dalam apa yang dapat dianggap dapat diterima secara klinis.



Karena masukan kami tentang Risiko Tambahan adalah perkiraan perhitungan, fungsi Keanggotaan mencerminkan bahwa dengan memiliki kurva jangkauan yang sangat besar.

Akhirnya, kita memiliki fungsi Keanggotaan output kita. Perhatikan bahwa domain variabel 'Kemungkinan' bias ke kanan - ini untuk menjelaskan fakta bahwa seseorang yang terpengaruh lebih mungkin secara umum untuk menunjukkan gejala daripada seseorang yang tidak.

Untuk basis aturan, kami membangun total sepuluh aturan yang mengatur Sistem Logika Fuzzy. Sementara sistem akan telah jauh lebih akurat dengan basis aturan yang lebih besar, kami tidak memiliki cukup Input atau Linguistik Variabel untuk menghasilkan basis aturan yang lebih besar. Basis aturan diberikan di bawah ini sebagai tabel:

