复旦大学
FUDAN UNIVERSITY
邮编（Zip Code）：200433/200438
电话（Tel.）：86-21-65642222
网址（Web.）：www.fudan.edu.cn

REVIEW OF THE FOREIGN SCIENTIFIC CONSULTANT on the dissertation work of Madina Suleimenova on the topic: «Machine learning in predicting the risk of developing cardiovascular disease» submitted for the degree of Doctor of Philosophy (PhD) in the educational program: 8D06114 – Artificial Intelligence in Medicine

The dissertation devoted to the development and validation of interpretable machine learning models for predicting cardiovascular disease (CVD) risk in adults aged 65 years and older addresses a highly relevant and practically important problem in modern medical informatics. Cardiovascular diseases remain among the leading causes of mortality and disability, and in older populations, the risk increases further due to age-related changes, multimorbidity, and lifestyle factors. In this context, improving early risk stratification and preventive decision-making is of particular importance.

The relevance of the dissertation topic is determined by the need for prediction tools specifically tailored to the 65+ population, for which the transportability of traditional risk scores may be limited. The work is further strengthened by the author's focus on interpretable models suitable for clinical implementation, as well as by the objective assessment of model quality and the clinical utility of decision thresholds. The study follows modern requirements for clinical prediction models, including a rigorous internal evaluation scheme, analysis of probabilistic performance and calibration, justification of decision thresholds, and verification of transportability on independent data.

The dissertation is well structured and logically coherent: its chapters consistently develop the research problem from the literature background and methodology to experimental results, discussion, and practical implementation. Chapter 1 provides a systematic review of national and international research on cardiovascular risk prediction in older adults, covering both classical risk assessment approaches and contemporary machine learning methods. The author clearly motivates the importance of interpretability, probability calibration, and clinical utility evaluation, which leads naturally to the study's aim and objectives.

Chapter 2 describes the materials and methods in detail. The author presents the cohort characteristics, data preprocessing procedures, principles of feature construction from clinical and laboratory indicators, and the selection and tuning of machine learning algorithms. A notable strength is the correct validation strategy: nested cross-validation is applied to separate hyperparameter optimization from independent model evaluation, thereby reducing the risk of optimistic bias. In addition, the dissertation evaluates probabilistic performance and calibration, which are

复旦大学
FUDAN UNIVERSITY
邮编 (Zip Code)：200433/200438
电话 (Tel.)：86-21-65642222
网址 (Web.)：www.fudan.edu.cn

essential for using the model as a risk calculator in practice.

Chapter 3 reports the main results. The author trained and compared model configurations on the development cohort and generated out-of-fold probabilities to construct ROC/PR curves, calibration plots, and subsequent clinical assessments. A key advantage is that the study does not rely solely on discrimination metrics but also reports probabilistic metrics and calibration analysis, aligning with contemporary standards for clinical prediction modeling. To support clinical applicability, the author used Decision Curve Analysis (DCA) to evaluate clinical benefit across thresholds and established a three-level risk scale (low/intermediate/high), enabling practical interpretation of predictions.

A particularly strong methodological aspect of the dissertation is the assessment of model transportability. The author conducted external validation on an independent cohort substantially larger than the development set, which allows a well-grounded evaluation of model performance beyond the training population. Furthermore, the dissertation addresses probability recalibration when transferring the model to an external population, improving the correctness of predicted probabilities and increasing the practical value of the results.

The doctoral student's publication activity deserves special mention. More than 10 scientific papers have been published on the topic of his dissertation research, including articles in journals indexed in Scopus and Web of Science, as well as publications in peer-reviewed journals recommended by the Committee for Quality Assurance in Science and Higher Education of the Republic of Kazakhstan. The author has received copyright certificates for results related to the software implementation and/or computational components of the developed risk assessment system. The presence of publications and confirmed intellectual property results demonstrates the high level of validation and practical significance of the work.

Overall, the dissertation is completed at a high scientific and methodological level, with well-supported conclusions and a clear practical orientation. The work demonstrates scientific novelty through the development of an interpretable risk stratification model for an older population, supported by objective internal validation, DCA-based threshold justification, and confirmation of transportability on an independent external cohort.

Based on the analysis of the obtained results, it can be concluded that the dissertation work of Madina Suleimenova is a completed scientific research project and can be recommended for defense for the degree of Doctor of Philosophy (PhD).

Sincerely,

Siming Chen, Associate Professor
School of Data Science, Fudan University
Website: http://simingchen.me Email: simingchen@fudan.edu.cn
Phone: +86-18801114625

8D06114 – «Медицина саласындағы жасанды интеллект» білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алу үшін ұсынылған Мадина Сулейменованың «Жүрек-қан тамырлары ауруларының даму қаупін болжаудағы машиналық оқыту» тақырыбындағы диссертациялық жұмысына

ШЕТЕЛДІК ҒЫЛЫМИ КЕҢЕСШІНІҢ ПІКІРІ

65 жас және одан үлкен ересектерде жүрек-қан тамырлары ауруларының (ЖҚА) қаупін болжау үшін түсіндірілетін машиналық оқыту модельдерін әзірлеу және валидациялауға арналған диссертация қазіргі заманғы медициналық информатикадағы өте өзекті және практикалық маңызды мәселені қарастырады. Жүрек-қан тамырлары аурулары өлім мен мүгедектіктің негізгі себептерінің бірі болып қала береді, ал егде жастағы адамдарда жасқа байланысты өзгерістерге, көп аурушаңдыққа және өмір салты факторларына байланысты қауіп одан әрі артады. Осыған байланысты ерте тәуекел стратификациясын және алдын алу шешімдерін қабылдауды жақсарту ерекше маңызды.

Диссертация тақырыбының өзектілігі 65+ жастағы популяцияға арнайы бейімделген болжау құралдарының қажеттілігімен анықталады, олар үшін дәстүрлі қауіп ұпайларының тасымалдануы шектеулі болуы мүмкін. Жұмыс автордың клиникалық енгізуге жарамды түсіндірілетін модельдерге баса назар аударуымен, сондай-ақ модель сапасын және шешім қабылдау табалдырықтарының клиникалық пайдалылығын объективті бағалаумен одан әрі күшейтіледі. Зерттеу клиникалық болжау модельдеріне қойылатын заманауи талаптарға, соның ішінде қатаң ішкі бағалау схемасына, ықтималдық көрсеткіштерін талдауға және калибрлеуге, шешім қабылдау шектерін негіздеуге және тәуелсіз деректер бойынша тасымалдануын тексеруге сәйкес келеді.

Диссертация жақсы құрылымдалған және логикалық тұрғыдан үйлесімді: оның тараулары зерттеу мәселесін әдебиеттер мен әдіснамадан бастап эксперименттік нәтижелерге, талқылауға және практикалық іске асыруға дейін үнемі дамытады. Бірінші тарауда егде жастағы адамдарда жүрек-қан тамырлары қаупін болжау бойынша ұлттық және халықаралық зерттеулерге жүйелі шолу жасалған, онда классикалық тәуекелді бағалау тәсілдері де, заманауи машиналық оқыту әдістері де қамтылған. Автор түсіндірілудің, ықтималдықты калибрлеудің және клиникалық пайдалылықты бағалаудың маңыздылығын айқын көрсетеді, бұл зерттеудің мақсаты мен міндеттеріне табиғи түрде әкеледі.

Үшінші тарауда зерттеудің негізгі нәтижелері берілген. Автор development-когортада әртүрлі модель конфигурацияларын оқытып, салыстырған, сондай-ақ ROC және PR қисықтарын, калибрлеу графиктерін және кейінгі клиникалық бағалауларды құру үшін out-of-fold ықтималдықтарын қалыптастырған. Жұмыстың маңызды артықшылығы — зерттеудің тек дискриминациялық метрикалармен шектелмей, ықтималдық

метрикалары мен калибрлеу талдауын да қамтуында. Бұл қазіргі клиникалық болжамдық модельдеу стандарттарына толық сәйкес келеді. Модельдің клиникалық қолданылуын негіздеу үшін автор Decision Curve Analysis (DCA) әдісін пайдаланып, әртүрлі шектердегі клиникалық пайдалықты бағалаған және практикалық түсіндіруге ыңғайлы үш деңгейлі тәуекел шкаласын — төмен, орташа және жоғары қауіп деңгейлерін — ұсынған.

Диссертацияның ерекше әдіснамалық аспектісі - модельдің тасымалдануын бағалау. Автор әзірлеу жиынтығынан айтарлықтай үлкен тәуелсіз когортада сыртқы валидация жүргізді, бұл оқыту популяциясынан тыс модельдің өнімділігін жақсы негізделген бағалауға мүмкіндік береді. Сонымен қатар, диссертация модельді сыртқы популяцияға ауыстырған кезде ықтималдықты қайта калибрлеуді қарастырады, болжамды ықтималдықтардың дұрыстығын жақсартады және нәтижелердің практикалық құндылығын арттырады.

Докторанттың басылымдық белсенділігі ерекше атап өтуге тұрарлық. Оның диссертациялық зерттеу тақырыбы бойынша 10-нан астам ғылыми мақала жарияланды, оның ішінде Scopus және Web of Science индекстелген журналдардағы мақалалар, сондай-ақ Қазақстан Республикасының Ғылым және жоғары білім беру саласындағы сапаны қамтамасыз ету комитеті ұсынған рецензияланған журналдардағы жарияланымдар бар. Автор бағдарламалық жасақтаманы енгізуге және/немесе әзірленген тәуекелдерді бағалау жүйесінің есептеу компоненттеріне қатысты нәтижелер үшін авторлық құқық сертификаттарын алды. Басылымдардың және расталған зияткерлік меншік нәтижелерінің болуы жұмыстың жоғары деңгейдегі валидациясы мен практикалық маңыздылығын көрсетеді.

Жалпы алғанда, диссертация жоғары ғылыми және әдіснамалық деңгейде, жақсы негізделген қорытындылармен және айқын практикалық бағдармен аяқталды. Жұмыс егде жастағы халық үшін түсіндірілетін тәуекел стратификациясының моделін әзірлеу арқылы ғылыми жаңалықты көрсетеді, оны объективті ішкі валидация, DCA негізінде шешім шектерін негіздеумен және тәуелсіз сыртқы когортада тасымалданғыштығын растаумен толықтырылған.

Алынған нәтижелерді талдау негізінде Мадина Сулейменованың диссертациялық жұмысы аяқталған ғылыми зерттеу жобасы болып табылады және философия докторы (PhD) дәрежесін алу үшін қорғауға ұсынылуы мүмкін деген қорытынды жасауға болады.

Құрметпен,

Симинг Чен, қауымдастырылған профессор
Фудань университеті
Деректер ғылымдары мектебі
Веб-сайт: http://simingchen.me
Электрондық пошта: simingchen@fudan.edu.cn
Телефон: +86-18801114625