CLEMSON

School of COMPUTING

Computer Science—Human-Centered Computing—Visual Computing

November 27, 2019

Re: Review of the PhD thesis dissertation entitled "Research and development of an information system for diagnosing diabetes based on Big Data technology tools" submitted for the degree of Doctor of Philosophy (PhD) majored in 6D070300 - "Information Systems"

To Whom It May Concern:

I am the C. Tycho Howle Director of the School of Computing at Cand am writing this letter in support of Assel Mukasheva in support of the defense of her dissertation entitle, "Research development of an information system for diagnosing diabetes based on Big Data technology tools."

The analysis of the current state of the issue of using BigData technologies in the field of medical services in the Republic of Kazakhstan showed that this technology is not widely used in the country. The capabilities of the developed system will reduce costs by diagnosing the disease at the initial stage, where treatment can be done without the use of expensive drugs, and subject to the usual diets and physical activity.

The problem solved in this thesis is of practical value for regulating the tasks that allow people to get a significant social effect due to the timely diagnosis of diabetes as well as maintaining a common knowledge base about this disease.

The significance and relevance of such studies lie in the fact that experimental studies have been carried out where methods of mathematical statistics and machine learning are considered to predict the growth of possible patients for years to come. Thus helps to pre-plan the costs of regional budgets for the purchase of insulin in the regions. In this regard, the dissertation work of doctoral student A.K. Mukasheva dedicated to the research and development of the information system of the support of endocrinologists for the diagnosis of diabetes mellitus and the prediction of the growth of patients with type II diabetes, based on BigData tools, is undoubtedly relevant for both practical and theoretical interests.

The most significant scientific results obtained by a doctoral student during research include:

1. The study investigated and established the set of acceptable solutions to statistical problems, and the results obtained made it possible to predict the possible number of patients for following years. The correlations between the population growth and patients with type II diabetes, as well as the increase in gross regional product and the growth of patients with diabetes are described. In addition, the relationship of population growth and gross domestic product with the growth of patients with diabetes in the Republic of Kazakhstan was determined.

2. The developed model of the processes of search, extraction, processing and analysis of data in large data arrays allows to store and process large volumes of structured and unstructured

data, and meanwhile, these data is protected from losses and hardware failures. The ecosystem created around Hadoop includes a number of other open source technologies that can complement and expand its core capabilities.

3. Conducted distributed computing, ensuring the effective implementation of the developed algorithms;

4. An information system for the diagnosis of diabetes based on BigData technology tools has been developed and tested. It was found that the system will facilitate the work of endocrinologists, as they can use not only their own knowledge, but also the global resources of the diabetes knowledge base.

The dissertation is aimed at solving an urgent problem of theoretical and applied value in the field of information systems. The theoretical significance of the work lies in the theoretical justification and comparative analysis of regression analysis methods and machine learning algorithm to predict the possible growth of patients in the Republic of Kazakhstan. Using a comprehensive set of statistics over the past 15 years to find patterns in statistics and build the necessary forecasting models. The practical significance of the work lies in the development of software products based on BigData technologies to improve the quality of early diagnosis of type II diabetes.

Proceeding from the above, PhD thesis dissertation entitled "Research and development of informational diabetes diagnosis system based on BigData technology tools" for its relevance, scientific novelty, is fully consistent with the rules of section 2 of the "Rules of awarding academic degrees," the Committee for Control of Education and Science of RK, and its author, Assel Mukasheva deserves the award of a Doctor of Philosophy (PhD) degree majored in 6D070300 - "Information Systems".

Sincerely,

Amy Apon, Ph.D.
C. Tycho Howle Director, School of Computing
Clemson University