

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК \_\_\_\_\_

Абрамчук Максим Александрович

Модели и методы сбора анкетных данных абитуриентов

Автореферат  
на соискание степени магистра технических наук  
по специальности 1-40 80 02 «Системный анализ, управление и обработка  
информации»

*(шифр и название специальности согласно учебному плану)*

---

*(подпись магистранта)*

Научный руководитель  
Никульшин Борис Викторович

*(фамилия, имя, отчество)*

кандидат технических наук

*(ученая степень, ученое звание)*

---

*(подпись научного руководителя)*

Минск 2019

## ВВЕДЕНИЕ

Беларусь входит в состав государств с большим образовательным резервом. Для того, чтобы образование соответствовало международному, ряд ВУЗов принял решение перейти на двухуровневую систему: базового и последипломного образования. В соответствии с Болонским процессом на замену специалитета пришли два уровня – бакалавриат и магистратура. Бакалавриат призван удовлетворять массовый спрос на высшее образование, магистратура — способствовать формированию профессиональной элиты. После данного перехода возрос спрос на магистерские программы. Важным этапом является эффективное управление процессом отбора абитуриентов при переходе с одного уровня на другой.

Требования к качеству знаний магистров продиктованы существующими профессиональными стандартами, в составлении которых участвует потенциальный работодатель. В связи с этим возникает необходимость к повышению результативности целевого обучения. При решении этой задачи одним из первых этапов является отбор абитуриентов приёмной комиссией. Поэтому от качества его организации во многом зависит тот результат, который покажут будущие магистры.

Задача отбора характеризуется значительным разнообразием параметров, оценивающих качество предшествующей подготовки, характеризующих опыт, возможности, способности и знания, которые накоплены и получены студентом за предшествующий период обучения в ВУЗе, и выражаемых количественными или качественными переменными.

Процесс поступления абитуриентов на магистерские программы в целом является более персонализированным, чем при поступлении на бакалаврские программы. Как правило, необходимо выявить некоторую предрасположенность будущего магистранта к научной работе, получить рекомендации, серьезное внимание уделить оценке личностных качеств кандидата на бюджетное место.

Сбор и обработка анкетных данных абитуриентов, поступающих в магистратуру – трудоёмкие процессы, ведь необходимо собрать и обработать данные с многочисленного числа студентов.

Для автоматизации процесса отбора анкетных данных были разработаны и обучены алгоритмы машинного обучения, которые выдают рекомендуемый статус для анкеты: будет ли она рекомендована для прохождения к следующему этапу отбора в магистратуру.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

**Актуальность** темы обусловлена большими трудозатратами и нерелевантными результатами обработки анкетных данных абитуриентов.

### **Цель и задачи исследования**

Целью магистерской работы является повышение качества работы приёмной комиссии за счет применения методики отбора абитуриентов в магистратуру по результатам вступительных испытаний и прогнозирования результатов обучения путём автоматизация процесса сбора анкетных данных в дистрибутиве Python Anaconda с использованием алгоритмов машинного обучения.

### **Основные положения, выносимые на защиту**

- 1) Разработанные алгоритмы машинного обучения для автоматизации процесса сбора анкетных данных.
- 2) Разработанная параметрическая анкета абитуриента.

### **Структура и объем работы**

Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Работа состоит из введения, трёх глав и заключения, списка использованных источников и приложений. Общий объем диссертации – 47 страниц. Работа содержит 3 таблицы, 16 рисунков. Библиографический список включает 19 наименований.

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

Во введении рассмотрено современное состояние проблемы обработки анкетных данных абитуриентов, определены основные направления исследований, а также дается обоснование актуальности темы диссертационной работы. Сформулированы ее цель и задачи, даны сведения об объекте исследования и обоснован его выбор, представлены положения, выносимые на защиту, а также, структура и объем диссертации.

Первый раздел является теоретической базой магистерской работы, в нем рассмотрены существующие системы сбора и обработки данных. Изучены системы искусственного интеллекта и программное обеспечение для данных систем.

Во втором разделе разработана параметрическая модель абитуриента, выявлены параметры, влияющие на эффективность отбора, а также предложена

методика вступительных испытаний на основании прогнозирования успешности обучения с использованием экспертных систем.

В третьем разделе раскрыт вопрос применения систем искусственного интеллекта при сборе и обработке данных, представлены разработанные классификаторы.

## **ЗАКЛЮЧЕНИЕ**

Результатом выполнения диссертационной работы является разработанный классификатор и алгоритмы машинного обучения, которые позволяют проанализировать заполненные анкеты абитуриентов согласно вероятности успешной защиты.

В ходе написания данной работы были поставлены и решены следующие задачи:

- рассмотрены теоретические основы систем сбора и обработки данных;
- изучены системы искусственного интеллекта и программное обеспечение систем ИИ;
- автоматизирован процесс сбора и обработки анкетных данных в дистрибутиве Python Anaconda с использованием алгоритмов машинного обучения;

На основе проведенных исследований в качестве метода для автоматизации процесса отбора анкетных данных рекомендованы к использованию 3 классификатора. После внедрения алгоритма существенно сократятся трудозатраты.

Хочется отметить, что проект подлежит масштабированию, так как ежегодно база анкет абитуриентов пополняется, появляются новые признаки, следовательно, классификаторы можно обучать на новых выборках.

Также для имплементации выбранного классификатора в дальнейшем может быть разработан web-сервис, с которым будет взаимодействовать модуль системы сбора и обработки анкетных данных посредством передачи JSON данных.