

**Развернутая аннотация магистерской программы
«Интеллектуальный анализ данных и биоинформатика»
направление подготовки 01.04.02 Прикладная математика и информатика**

Руководитель программы: директор НОЦ компьютерных наук и технологий, доктор технических наук, доцент Замятин А.В.

1. Цель создания магистерской программы

Магистерская программа «Интеллектуальный анализ данных и биоинформатика» является уникальной пилотной 2-х летней образовательной программой, разработанной в рамках «дорожной карты» развития Томского государственного университета. Важной отличительной особенностью этой новой программы является ее выраженная междисциплинарная направленность на компьютерные науки, информационные технологии, прикладную математику, а также прикладные аспекты биоинформатики – сложной и перспективной области человеческой деятельности на стыке компьютерных наук, биологии, медицины.

Цель программы – подготовка высококвалифицированных глобально конкурентоспособных специалистов, компетентных не только в области интеллектуального анализа данных, но и способных к применению соответствующих методов к области анализа биологических, биомедицинских данных и извлечения знаний.

2. Концепция магистерской программы

Стремительная технологическая эволюция последних лет в сфере информационно-коммуникационных технологий позволила сформировать существенный задел в части развитой программно-аппаратной инфраструктуры, поддерживающей накопление и постоянное пополнение архивов данных различной природы и назначения. Однако, как извлечь полезные знания из таких архивов, содержащих многомерные, разнородные, неполные, неточные, противоречивые данные? Как это сделать эффективно, если объемы данных измеряются гигабайтами и терабайтами?

Ответы на эти и многие другие вопросы дает активно развивающаяся область интеллектуального анализа данных, связанная с высокопроизводительной аналитической обработкой данных, и направленная на то, чтобы оперативно извлекать из значительных массивов накопленных и поступающих данных ценные экспертные знания, поддерживая эффективную управленческую деятельность.

Особый интерес сегодня представляют специалисты в области компьютерных наук и информационных технологий, способные после соответствующей подготовки решать задачи анализа данных в сложной и перспективной области человеческой деятельности на стыке компьютерных наук, биологии, медицины – биоинформатики.

Всё это определяет содержание магистерской программы, предполагающее формирование компетенций в следующих областях:

- разработка новых и модернизация существующих алгоритмов и информационно-программных комплексов для эффективного интеллектуального анализа и интерпретации данных;

- информационные технологии интеллектуального анализа данных, включая алгоритмы и структуры данных, применяемые для анализа больших массивов данных, которые формируются, в том числе с помощью современных биотехнологических и биомедицинских методов;

- прикладные аспекты состояний биообъектов, структурной и молекулярной биологии, геномики, первичной обработки и хранения экспериментальных банков медико-биологических данных.

В концепции подготовки магистров учтены не только вопросы интеллектуального анализа данных и биоинформатики, но и результаты научно-исследовательского задела ТГУ, базирующиеся на применении информационных технологий и компьютерных наук в биологии и медицине.

3. Обоснование потребности в магистрах данного профиля

В конце XX века в биологии и биомедицине начали активно применять вычислительную технику, количество разнородных экспериментальных данных стало неуклонно и интенсивно возрастать, существенно превышая возможности их анализа человеком. Вместе с тем, интенсивное развитие компьютерных наук, методов интеллектуального анализа данных, прикладных аспектов математической статистики и теории вероятностей применительно к задачам интеллектуальной обработки биологических и биомедицинских данных позволили бурно развивать новую область науки – биоинформатику.

Основными задачами биоинформатики считают:

- первичная обработка и хранение экспериментальных данных;
- дружественный интерфейс для удаленного оперативного доступа к хранимым данным;
- интеллектуальный анализ медико-биологических данных;
- поддержка процессов ускоренного конструирования лекарственных средств.

Учитывая стратегическую важность биоинформатики, эта область активно развивается в США, Австралии, странах Европы и Азии. Подключилась к этим процессам и Россия – различные по задачам, уровню инфраструктурного и кадрового обеспечения подразделения биоинформатики появились в Москве, Санкт-Петербурге, Новосибирске, Нижнем Новгороде.

4. Условия обучения

Срок обучения – 2 года (4 семестра)

Форма обучения – очная.

5. Набор студентов и требования к поступающим в магистратуру

Приём в магистратуру осуществляется на конкурсной основе по результатам вступительных испытаний на уровне бакалавриата по направлению «Математика», «Математика и компьютерные науки», «Фундаментальная информатика и информационные технологии», «Прикладная математика и информатика», «Математическое обеспечение и администрирование информационных систем», «Механика и математическое моделирование», «Информатика и вычислительная

техника», «Информационные системы и технологии», «Прикладная информатика», «Программная инженерия», «Прикладная математика» и др.

Поступающие в магистратуру, имеющие диплом бакалавра по указанным направлениям, проходят вступительные испытания в форме собеседования с целью установления у поступающего наличия следующих компетенций:

- 1) способность практического использования профессиональных знаний в области прикладной математики и математической статистики, компьютерных наук, информационных технологий;
- 2) способность проводить научные исследования и получать новые научные и прикладные результаты;
- 3) способность пользоваться русским и иностранным языками как средством делового общения и социальной мобильности;
- 4) обоснование целей обучения в магистратуре и мотивация.

6. Кадровая, методическая и материально-техническая базы магистерской программы

Реализация ООП магистратуры «Интеллектуальный анализ данных и биоинформатика» обеспечивается научно-педагогическими кадрами, имеющими базовое образование, соответствующее профилю преподаваемой дисциплины, и ученую степень или опыт деятельности в соответствующей профессиональной сфере и систематически занимающимися научной и (или) научно-методической деятельностью. К образовательному процессу по дисциплинам профессионального цикла привлечены не менее 20 процентов преподавателей из числа действующих руководителей и ведущих работников профильных организаций, предприятий и учреждений. Не менее 80 процентов преподавателей (в приведенных к целочисленным значениям ставок), обеспечивающих учебный процесс по профессиональному циклу и научно-исследовательскому семинару, имеют ученые степени и ученые звания, при этом ученые степени доктора наук или ученое звание профессора имеют не менее 12 процентов преподавателей.

Кадровый состав постоянно усиливается привлечением к образовательному процессу (чтение лекций, семинаров, аттестационные комиссии и т.п.) ведущих отечественных и зарубежных ученых и специалистов предприятий (ООО «F5», ООО «ИндорСофт», ООО «Интант» и др.), созданными ранее с учетом имеющегося научно-технического задела ТГУ в ИТ сфере при решающем участии его сотрудников, а также университетов партнеров (Сибирский государственный медицинский университет, Goldsmiths College (Великобритания) и др.).

7. Содержание программы и общая характеристика учебного плана

ОСНОВНЫЕ ДИСЦИПЛИНЫ:

- Введение в интеллектуальный анализ данных
- Многомерный анализ и визуализация данных
- Прикладные аспекты машинного обучения
- Сбор, хранение и обработка данных в гетерогенных распределенных компьютерных сетях
- Технологии высокопроизводительной обработки данных
- Облачные вычисления и технологии

- Скриптовые языки
- Алгоритмы и структуры данных
- Статистический анализ последовательностей
- Обработка естественного языка
- Математические модели систем поддержки принятия решений
- Введение в биоинформатику
- Биология клетки и биохимия
- Молекулярная генетика
- Основы биотехнологии
- Молекулярная биология и анализ молекулярных последовательностей
- Сравнительная геномика
- Специализированные биологические базы данных
- Транскриптомный анализ
- Управление проектами по разработке программного обеспечения

В план обучения магистров обязательно входит научно-исследовательская работа.

Цель НИР:

- приобретение практических навыков и компетенций научно-исследовательской деятельности,
- закрепление и углубление теоретической подготовки магистра.

НИР выполняется на предприятиях, НИИ, подразделениях ТГУ, ведущих исследования и разработки в предметной области ООП по заказу предприятий и организаций. Место проведения выбирается индивидуально для каждого магистра и утверждается выпускающим подразделением университета.

8. Перспективы научно-исследовательской деятельности в связи с развитием ТГУ и потребностями Томского региона

Томский государственный университет обладает развитой научно-педагогической школой в области компьютерных наук, информационных технологий и прикладной математической статистики. В настоящее время она включает такие направления, как вычислительная геометрия и машинная графика (доктор технических наук, профессор Ю.Л. Костюк), геоинформационные системы и технологии (доктор технических наук А.И. Рюмкин и доктор технических наук, профессор А.В. Скворцов), информатика динамических систем (доктор технических наук, лауреат Государственной премии, профессор В.В. Поддубный), программная инженерия (доктор физико-математических наук, профессор О.А. Змеев), информационные технологии в экономике и управлении (кандидат физико-математических наук, профессор Б.А. Гладких), анализ и оптимизация распределенных информационных процессов (доктор технических наук, профессор С.П. Сущенко).

Перспективы построения в университете Центра трансляционной медицины и Центра биоинформатики, диктуемые потребностями в современных биотехнологических и биомедицинских методах с увеличивающейся в них ролью информационных технологий эффективной обработки данных также обуславливают необходимость разработки и реализации образовательных программ, позволяющих ликвидировать острую нехватку уникальных междисциплинарных специалистов в этих областях.

9. Перспективы профессиональной деятельности и трудоустройства

Выпускники магистерской программы «Интеллектуальный анализ данных и биоинформатика» готовы к профессиональной деятельности по интеллектуальному анализу данных различной природы и назначения, включая области, связанные с биоинформатикой - в научно-исследовательских институтах, специализированных лабораториях и биотехнологических компаниях.

В России и за рубежом есть потенциальный круг организаций, которым требуются специалисты с развитыми компетенциями в области биоинформатики: научно-исследовательские институты и университеты, фармакологические и биотехнологические компании, осуществляющие свою деятельность в молекулярной биологии, молекулярной генетике, генетической инженерии, биотехнологии, конструировании лекарственных средств, клеточной биологии.