

## ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ И ВЕРОЯТНОСТНЫХ МЕТОДОВ ОБРАБОТКИ ДАННЫХ ТЕХНОЛОГИЧЕСКОГО НАЗНАЧЕНИЯ С УСКОРОЕНИЕМ НА GPU

**Добрынин В.Ю.** (Университет ИТМО, Санкт-Петербург)  
**Научный руководитель – к.т.н., доцент Сисюков А.Н.**  
(Университет ИТМО, Санкт-Петербург)

В работе рассмотрено применение статистических и вероятностных методов обработки данных технологического назначения с ускорением на GPU. Рассмотрены методы регрессии и классификации на основе алгоритмов повышения градиента и их применимость для обработки на ускорителях. Выполнен кластерный анализ на GPGPU для предварительной сегментации набора данных с помощью алгоритма пространственной кластеризации, основанной на плотности слоя данных с шумами. В качестве исходных данных взяты данные из открытых источников и стенд на базе процессоров NVIDIA GP102.

### **Введение.**

В процессе современного производства конкурентной продукции заимствуются методы автоматизации не только механического труда человека, но и интеллектуальной его деятельности. В последнем случае используются методы и подходы, связанные с понятием искусственного интеллекта. Машинное и глубокое обучение, как совокупность алгоритмов и методологий, приближающих результат вычислительного процесса к результатам интеллектуальной деятельности человека, становится неотъемлемой частью процессов производственной автоматизации.

### **Основная часть.**

Большой объем информации, хранимой в структурированном и неструктурированном виде и получаемой с IoT устройств, смарт-оборудования и компонентов киберфизических систем, требует обработки в целях извлечения значимой информации для рекомендательных систем, систем оповещения и принятия решений. Алгоритмы классификации, кластеризации и предиктивного анализа играют существенную роль в интерпретации входящих в подсистему обработки производственных данных.

Современные сетевые кластерные системы на базе последних поколений многоядерных центральных процессоров (ЦПУ) способны справляться с обработкой больших потоков информации. Зачастую время обработки даже при многопроцессорных платформах, многоядерности и гиперпоточности неприемлемо для переобучения и инференса интеллектуальной модели в реальном времени. Расширение вычислительного кластера добавлением новых нод повысит общую производительность, и в то же время уменьшит энергоэффективность и экономическую привлекательность системы обработки данных. Использование аппаратных ускорителей (GPU) альтернативно позволяет повысить производительность, а в случае глубокого обучения и больших объемов данных дает единственно приемлемое время обработки.

В текущей работе представлен подход к обработке неструктурированной информации, представляемой в виде текстовых документов различной тематики и связанной с производственным процессом. Задачей является задействовать методы классификации логов, получаемых со смарт-оборудования или компонентов киберфизических систем, по признакам критичности информации, определяющей возможность фатального сбоя, и общего содержания с извлечением режимов работы оборудования и возможностью применения регрессионного анализа на данных, получаемых в дополнение к структурированным. Подход предлагается также применять для анализа производственной сопроводительной

документации по признакам финансового, логистического, конструкторского и технологического наполнения с автоматическим отнесением документа к определенному типу и поиска подобий в документах.

Предобработка текстовых логов на основе количественных метрик, таких как объем числовых, текстовых и специфических для оборудования тегированных лейблов, включая общее число символов, предложена на основе кластерного анализа с алгоритмом пространственной кластеризации, основанной на плотности слоя данных с шумами. Получаемые с оборудования логированные данные предварительно обрабатываются и кластеризуются на основе предложенных метрик. Результат кластеризации позволяет оценить наличие групп, схожих по метрикам.

Классификация логов по признакам критичности информации требует наличия обучающих выборок в отличии от этапа кластеризации данных. Подготовка подобных выборок является отдельной задачей. Обучение и инференс модели данных предложено проводить на базе алгоритма градиентного бустинга, лучше подходящего для задач обработки с гетерогенными данными в сравнении с нейронными сетями, где более подходяще однородные данные.

Предикативную составляющую относительно работы оборудования вносит регрессионный анализ на основе алгоритмов градиентного бустинга с деревьями решений.

### **Выводы.**

В ходе работы были исследованы GPU ориентированные реализации рассмотренных алгоритмов. Обработка данных полученных из открытых источников проводилась с использованием стенда с тремя ускорителями 1080TI, общей производительностью 34 TFLOPS и памятью 33 Gb на операционной системе Ubuntu Bionic Beaver.

Добрынин В.Ю. (автор)

Подпись

Сисюков А.Н. (научный руководитель)

Подпись