

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

УДК 004.85:811

Войтиховский
Павел Дмитриевич

«Адаптивные средства обучения иностранным языкам»

Автореферат

на соискание степени магистра технических наук
по специальности 1-40 80 06 «Искусственный интеллект»

Научный руководитель
Давыденко И. Т.
к.т.н., доцент,
доцент кафедры ИИТ

Минск 2021

КРАТКОЕ ВВЕДЕНИЕ

В современном мире знание иностранных языков является необходимостью, которая обуславливается различными факторами. В первую очередь, английский язык является стандартом международной коммуникации, что является причиной использования этого языка на конференциях, в профессиональной коммуникации, в написании профессиональной литературы и научно-исследовательских работ. Второй причиной для изучения иностранных языков является то, что появляется возможность чтения художественной и технической литературы в оригинале. Это особенно актуально для людей, работающих с быстро развивающимися технологиями, когда ожидание перевода документации стоит очень дорого.

Процесс изучения иностранных языков является довольно длительным, так как он затрагивает такие аспекты, как изучение лексики, грамматики, произношения, а также тренировку восприятия иностранной речи и способов самовыражения на иностранном языке. В то же время у людей зачастую нет необходимого количества времени для того, чтобы качественно проработать каждый из этих аспектов. В результате этого появляются различные проблемы, такие как языковой барьер, когда человеку тяжело выразить свои мысли, отсутствие понимания различных диалектов языка или непонимание культурных особенностей, которые ведут за собой к непониманию различных фраз. В связи с этим, критически важной задачей является сокращение времени, затрачиваемого на каждый из аспектов изучения языка.

Существует множество различных причин, которые влияют на время и качество изучения языка. К таким причинам можно отнести то, что пользователям необходимо самостоятельно искать всю необходимую для изучения информацию. Время, затрачиваемое на поиск информации увеличивается с каждым днем, так как ее количество ежедневно растет, а автоматизированных способов выделять релевантную информацию нет, поэтому эта задача возлагается на плечи обучающихся. Следовательно, актуальной является задача разработки средств обучения иностранному языку, которые бы позволили сократить время подбора материалов и ускорить процесс обучения при самостоятельном обучении.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследователя

Целью данной работы является сокращение затрачиваемого времени на изучение иностранного языка за счет внедрения адаптивного подхода в обучающие системы. Данный подход основан на использовании рекомендательной модели, которая позволяет генерировать наборы слов для изучения, базируясь на индивидуальных характеристиках пользователя, тем самым, уменьшая затрачиваемое время на поиск релевантных материалов к изучению.

Для достижения поставленной цели научного исследования были решены следующие задачи:

- проанализированы существующие подходы изучения иностранных слов;
- проанализированы способы реализации рекомендательных систем;
- проанализированы подходы к разработке мобильных приложений, предоставляющих удобный пользовательский интерфейс;
- сформулированы требования, предъявляемые адаптивным системам изучения иностранного языка;
- разработана и реализована рекомендательная модель, которая рекомендует слова для изучения на основе интересов пользователя, а также его уровня владения иностранным языком и способностями к его изучению;
- оценено качество реализованной модели;
- спроектирована и реализована адаптивная система изучения иностранного языка на основе выдвинутых требований;
- протестирована работа адаптивной системы изучения иностранного языка.

Объектом научного исследования являются адаптивные системы изучения иностранных языков.

Предметом научного исследования являются методы и средства разработки адаптивных систем изучения иностранных языков.

Новизна полученных результатов

Была спроектирована и разработана рекомендательная модель изучения иностранных слов, которая генерирует пользователю множество слов для изучения. Данная модель генерирует множество слов, основываясь на индивидуальных характеристиках пользователя, его интересах, мотивации

и уровню знания иностранного языка. Для реализации данной модели был спроектирован алгоритм, для реализации которого был использован кластерный анализ. Далее данная рекомендательная модель была интегрирована в адаптивную систему изучения иностранного языка, что позволило уменьшить затрачиваемое на подбор и изучение слов время пользователя.

Положения, выносимые на защиту

- рекомендательная модель изучения иностранного языка, реализованная на основе методов кластерного анализа и предназначенная для генерации к изучению множества иностранных слов, основываясь на индивидуальных характеристиках пользователя, его мотивации и интересах;
- адаптивная система для изучения иностранного языка, разработанная с использованием предложенной рекомендательной модели. Данная система реализована в виде мобильного приложения для изучения английского языка.

Апробация результатов

Результаты диссертационной работы обсуждались на конференции «Информационные технологии и системы 2020» (БГУИР, Минск, Беларусь, 2020).

Опубликованность результатов исследования

По теме диссертации опубликована 1 печатная работа в сборнике материалов международной конференции.

Структура и объем диссертации

Диссертация состоит из введения, трех глав, заключения, списка использованных источников и списка публикаций автора. Общий объем работы составляет 73 страницы, из которых основного текста 66 страниц, 47 рисунков, список использованных источников состоит из 28 наименований на 2 страницах.

КРАТКОЕ СОДЕРЖАНИЕ

Во **введении** обоснована актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначены актуальные задачи, решению которых посвящена диссертационная работа, поставлена цель и задачи исследования.

В **первой главе** были рассмотрены причины, которые мотивируют пользователей изучать иностранные языки. Описаны способы изучения иностранного языка и приведены достоинства и недостатки каждого из методов, а также даны ссылки на ресурсы, которые можно использовать для изучения тем или иным способом. В частности были описаны методы изучения лексики, приведены проблемы, с которыми сталкиваются пользователи, и описаны преимущества и недостатки каждого из способов изучения лексики. Также была аргументирована задача разработки адаптивной системы изучения иностранного языка и определены требования к данной системе:

- Эффективность. Эффективная система обучения иностранного языка – это система, которая будет помогать пользователю добиться необходимых результатов за кратчайшие сроки.

- Простота. Это требование, которое предъявляется адаптивной системе изучения иностранного для того, чтобы упростить пользовательский опыт и тем самым увеличить эффективность изучения языка.

- Гибкость. Под гибкостью понимается возможность системы предоставить пользователю выбор сроков обучения, темпа, форм и методов обучения.

Во **второй главе** была разработана и описана рекомендательная модель. Данная модель предназначена для генерации набора таких слов, которые будут интересны пользователю, подходить его уровню английского языка. Она построена на основе кластерного анализа и работает следующим образом:

- Шаг 1: с помощью метода k -средних множество пользователей разбивается на k непересекающихся кластеров, где каждый пользователь может состоять только в 1 кластере.

- Шаг 2: определяется кластер, к которому будет относиться пользователь, которому необходимо сгенерировать набор слов для изучения.

- Шаг 3: формируется список слов, которые изучают пользователи в этом кластере.

- Шаг 4: из этого списка слов удаляются слова, которые уже есть у

пользователя, для которого генерируется набор слов для изучения.

Также были описаны индивидуальные характеристики пользователя, на которых будет строиться рекомендательная модель и то, как они будут представлены: данные представляются в виде двумерного массива, в котором каждая строка описывает пользователя, а каждый столбец – это некоторая характеристика пользователя. К таким характеристикам были отнесены:

- возраст;
- время, проводимое в приложении;
- количество изученных слов;
- количество слов на изучении;
- количество слов из категории «одежда»;
- количество слов из категории «мода»;
- количество слов из категории «семья»;
- количество слов из категории «неправильные глаголы»;
- количество слов из категории «игры».

В качестве способа нормализации был выбран метод минимакс (см. формулу 1).

$$y = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (1)$$

Для того, чтобы оценить рекомендательную модель, использовался следующий подход: для каждого пользователя из тестовой выборки определяется кластер, к которому он относится. Далее формируется множество слов для изучения иностранного языка. Это множество называется рекомендациями для обучения. Множество слов, которые пользователь добавил в свой личный словарь называется добавленными словами. Для того, чтобы оценить рекомендательную модель, было найдено пересечение этих двух множеств и мощность результирующего множества разделена на мощность множества рекомендованных слов для обучения (см. формулу 2).

$$q = \frac{\sum_{i=1..n} |R_i \cap W_i|}{\sum_{i=1..n} |R_i|} \text{ где,} \quad (2)$$

R – множество рекомендованных слов,

W – множество добавленных пользователем слов.

В **третьей главе** были описаны переменные, которые используются в рамках рекомендательной модели, приведены гистограммы распределения этих переменных и описаны основные характеристики исходной выборки. Исходная выборка состоит из 500 человек, из которых большинство

представлено в возрасте от 20 до 40 лет, самому младшему пользователю 7 лет, а самому старшему – 76 лет. В данной выборке 264 человека мужского пола и 236 человек женского. Также было определено сколько времени тратят пользователи на изучения иностранного языка:

- есть пользователи, которые провели в приложении не больше 5 минут и затем удалили его;
- есть пользователи, которые ежедневно проводят в этом приложении больше 10 минут в день;
- наибольшее количество пользователей проводит в приложении от 40 до 270 минут в месяц.

Было определено сколько слов в месяц добавляют пользователи и в каких категориях:

- есть пользователи, которые не добавляют ни одного слова;
- есть пользователи, которые добавляют более 250 слов;
- наибольшее количество пользователей добавляет от 150 до 230 слов в месяц.

Также были приведены графики зависимости переменных друг от друга и сделаны некоторые выводы исходя из этих графиков. Основным выводом стало то, что пользователи в возрасте от 20 до 40 лет проводят в мобильном приложении больше времени, добавляют больше слов и их интересы более разносторонние, в то время, как пользователи до 20 лет проводят в приложении в среднем меньше времени и интересуются компьютерами и играми, а пользователи старше 35 лет интересуются словами из повседневной жизни: слова из категории «семья», «одежда» и т.д.

Для того, чтобы применить методы кластеризации, была выполнена нормализация данных. Для нормализации была выбрана минимакс нормализация, которая приводит числа в диапазон от 0 до 1.

Количество кластеров, которые будут взяты для нашей выборки не может быть случайным. Каждый кластер формируется исходя из расчета и сравнения расстояний данных. Поэтому был рассчитан «Within-Cluster-Sum-of-Squares», который позволил определить количество кластеров: 6 кластеров. В качестве меры расстояния был использован квадрат Евклидова расстояния.

После обучения рекомендательной модели была рассчитана ошибка по формуле 2 и она составила $q = 0.59$ или 59%.

После реализации рекомендательной модели была спроектирована и описана архитектура адаптивной системы изучения иностранного языка, приведены основные варианты использования данной системы, к которым

можно отнести:

- авторизацию и регистрацию;
- добавление слов для изучения и их перевод, примеры использования, синонимы и антонимы, а также категорию слова;
- удаление слов из списка слов для изучения;
- просмотр добавленных слов и рекомендуемых слов для изучения;
- изучение слов.

Также были описаны ошибки, которые могут возникать во время работы с этой системой и способы их устранения. Наибольшее количество ошибок пользователь может встретить при авторизации и регистрации, они связаны с неправильным вводом логина или пароля.

Библиотека БГУИР

ЗАКЛЮЧЕНИЕ

Иностранные языки – неотъемлемая часть жизни многих людей. С каждым годом все больше людей начинают изучать какие-то языки по тем или иным причинам. Изучение языка – это долгий и кропотливый труд, который требует хорошей дисциплины, систематичности, правильного подбора материалов для изучения и многого другого. Положение усложняет также то, что люди в большинстве своем не готовы выделять много времени на изучение языка. Поэтому актуальной становится задача подбора материалов, учитывающих уровень знания иностранного языка пользователем, его интересы, его способности к обучению, мотивацию и какие-то внешние факторы, влияющие на изучение языка.

Для решения этой задачи была спроектирована и реализована адаптивная система изучения иностранного языка в виде мобильного приложения. Данная система предоставляет пользователю широкий спектр возможностей, который включает в себя следующее: изучение иностранных слов на основе интервального алгоритма, добавление для изучения иностранных слов, просмотр списка добавленных слов, а также рекомендаций.

Для того, чтобы составить рекомендации слов, была спроектирована и реализована рекомендательная модель и разработан алгоритм для оценки качества. Данная модель с помощью методов математической статистики и, в частности, кластерного анализа формирует список слов для изучения пользователем. Данный список слов формируется с учетом интересов пользователя, его уровня знаний иностранного языка и способностей к обучаемости.

В дальнейшей планируется повысить качество рекомендательной модели в аспекте увеличения релевантности материалов для изучения. Также планируется расширить рекомендательную модель до рекомендации не только слов, но и грамматических упражнений, и упражнений на восприятие речи. А также планируется развитие адаптивной системы изучения иностранного языка путем добавления возможности выполнять грамматические упражнения и упражнения на восприятие речи, планируется добавить возможность отслеживания прогресса изучения, возможность добавлять цели и устанавливать напоминания о занятиях.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Войтиховский, П. Д. Адаптивная система обучения английскому языку на основе кластерного анализа / П. Д. Войтиховский, К. А. Сафоновко, И. Т. Давыденко // Информационные технологии и системы 2020 (ИТС 2020) = Information Technologies and Systems 2020 (ITS 2020) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 18 ноября 2020 года) / редкол. : Л. Ю. Шилин [и др.]. – Минск : БГУИР, 2020. – С. 68 - 69.

Библиотека БГУИР