



Тема:

Искусственный интеллект для определения Перечня возможных профессий по заданным умениям.

Автор: Берлина Дарья Дмитриевна

Ученик 10 класса Г

Школа: ГБОУ №1799

Руководитель: Юрченков Иван Александрович

Преподаватель ДТ «Альтаир»



Актуальность проекта:

1. Экономия времени:

ИИ может помочь сэкономить время на выборе профессии.

2. Помощь в выборе на основе навыков и умений:

Это позволяет людям более эффективно использовать свои способности и достигать успеха в карьере.

3. Упрощение процесса выбора профессии для людей, которые ищут новую сферу деятельности.



Цель работы:

Создание системы, которая на основе анализа умений и навыков, взятых из резюме человека, сможет с большей точностью предложить ему наиболее подходящие профессии.





Задачи работы:

1. Сбор и анализ информации о существующих профессиях, требованиях к ним и необходимых умений.
2. Разработка алгоритмов и моделей для определения возможных профессий.
3. Оценка результатов работы системы и определение её эффективности для помощи пользователям в выборе профессии.



Методика выполнения работы:

Разработка алгоритма классификации текстов

Деревья решений

Тестирование на реальных данных





Используемые материалы

1. Использовано 4 датасета, найденных на Kaggle.
2. Использовано 48 классов профессий.

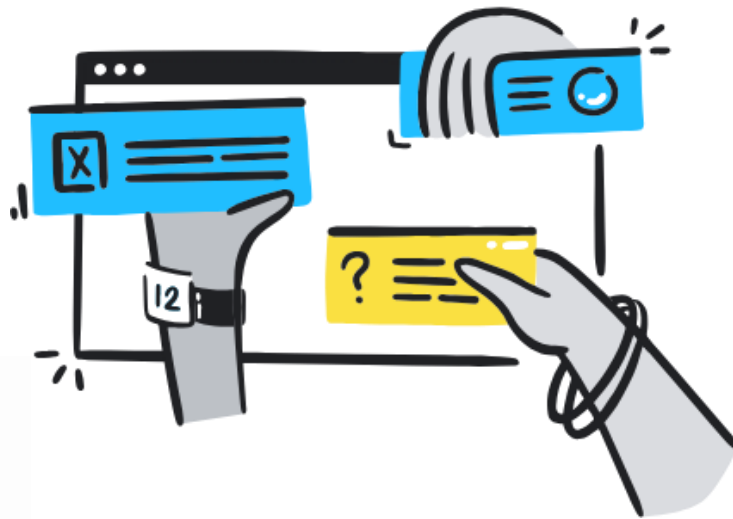


resume_dataset.csv (509.76 kB)							
Detail		Compact		Column			
▲ Category		☰		▲ Resume		☰	
Java Developer		8%		Education Details J...		1%	
HR		7%		Technical Skills Web...		1%	
Other (144)		85%		Other (165)		98%	
Data Science				Skills * Programming Languages: Python (pandas, numpy, scipy, scikit-learn, matplotlib), Sql, Java, ...			
Data Science				Education Details May 2013 to May 2017 B.E UIT-RGPVData ScientistData Scientist - MatelabsSkill ...			
Data Science				Areas of Interest Deep Learning, Control System Design, Programming in-Python, Electric Machinery, W...			



Используемое оборудование

- 1) Компьютер с доступом к интернету.
- 2) Программное обеспечение.
- 3) Использован графический процессор на Kaggle для ускорения классификации.





Полученные результаты:

Arts, "Education Details

August 2018 to January 2021 Intermediate Maths Mumbai,
Maharashtra Sunbeam academy , samne ghat , varanasi
Martial arts (fitness job)

Skill Details
Company Details

company - Sports Authority

description - I am 2nd dan black belt in karate (martial arts)

I am in a searching of personal trainer job for fitness.

I won 3 gold medals in national karate championship.

I won 7 gold medals in state karate championship.

3 times best player of the year of uttar pradesh award ..

Represented india and Selected for world karate championship

held at Croatia , Europe."

```
('Arts', 0.9714601108098396),  
( 'Health and fitness', 0.009714542899173033),  
( 'Advocate', 0.002319437758633723),  
( 'Testing', 0.0015235368925848511),  
( 'Java Developer', 0.0015109225054716557),  
( 'Sales', 0.0012280822294289011),  
( 'ARTS', 0.001048019135509105),  
( 'Mechanical Engineer', 0.0008508793678778857),  
( 'HR', 0.0008077588309203745),  
( 'DotNet Developer', 0.00075610928550518),
```




Полученные результаты:

Health and fitness, "SKILLS: Can work in Ms-Office (Word, Excel, and PowerPoint). Can easily operate internet & Gym Management Software (website & mobile application). Designing and promoting activities to meet customer demand and generate revenue;
Advertising and promoting the club or Centre to increase usage, considering market research;
Recruiting, training and supervising staff, including managing staff rotas;
Carrying out health and safety checks on the equipment and site;
Retaining existing members by providing good quality of service and by providing good sales programs;
Providing training to the sales staff to Deal with enquiries, complaints and emergencies
Delivering some fitness training or coaching in sports activities - often a good way of maintaining contact with customers; Preparing and checking budgets and generating revenue; Cashing-up and keeping stock records of supplements;
Utilizing Group-X studio by scheduling exciting grouper's batches; Keeping statistical and financial records;
Undertaking administrative tasks;
Promoting and marketing the business;"

```
('Health and fitness', 0.9655375379250037),  
( 'Sales', 0.00933290241326125),  
( 'FITNESS', 0.0032165408529075193),  
( 'CONSULTANT', 0.002291003193398095),  
( 'CHEF', 0.0013578530071138247),  
( 'Civil Engineer', 0.0012100703123382969),  
( 'SALES', 0.0010644930567410193),  
( 'HR', 0.0009873471829652512),  
( 'Arts', 0.000784144694887468),  
( 'ADVOCATE', 0.0007323846552751195),
```



Итоговая метрика

	precision	recall	f1-score	support
ACCOUNTANT	0.83	1.00	0.91	30
ADVOCATE	0.40	0.53	0.46	30
AGRICULTURE	0.43	0.19	0.26	16
APPAREL	0.53	0.38	0.44	24
ARTS	0.62	0.50	0.55	26
AUTOMOBILE	0.00	0.00	0.00	9
AVIATION	0.90	0.90	0.90	29
Advocate	1.00	1.00	1.00	12
Arts	1.00	1.00	1.00	19
Automation Testing	1.00	1.00	1.00	15
BANKING	0.65	0.52	0.58	29
BPO	0.00	0.00	0.00	5
BUSINESS-DEVELOPMENT	0.70	0.77	0.73	30
Blockchain	1.00	1.00	1.00	21
Business Analyst	1.00	1.00	1.00	15
CHEF	0.96	0.80	0.87	30
CONSTRUCTION	0.75	0.86	0.80	28
CONSULTANT	0.62	0.86	0.72	29
Civil Engineer	1.00	1.00	1.00	13
DESIGNER	0.86	0.93	0.89	27
DIGITAL-MEDIA	0.75	0.62	0.68	24
Data Science	1.00	1.00	1.00	22
Database	1.00	1.00	1.00	19
DevOps Engineer	1.00	1.00	1.00	29
DotNet Developer	1.00	1.00	1.00	16
ENGINEERING	0.89	0.80	0.84	30
ETL Developer	1.00	1.00	1.00	21
Electrical Engineering	1.00	1.00	1.00	16

	precision	recall	f1-score	support
FINANCE	0.78	0.93	0.85	30
FITNESS	0.86	0.83	0.84	29
HEALTHCARE	0.79	0.76	0.77	29
HR	0.96	0.92	0.94	52
Hadoop	1.00	1.00	1.00	23
Health and fitness	1.00	1.00	1.00	16
INFORMATION-TECHNOLOGY	0.68	0.83	0.75	30
Java Developer	1.00	1.00	1.00	46
Mechanical Engineer	1.00	1.00	1.00	21
Network Security Engineer	1.00	1.00	1.00	14
Operations Manager	1.00	1.00	1.00	21
PMO	1.00	1.00	1.00	16
PUBLIC-RELATIONS	0.75	0.64	0.69	28
Python Developer	1.00	1.00	1.00	26
SALES	0.46	0.55	0.50	29
SAP Developer	1.00	1.00	1.00	13
Sales	1.00	1.00	1.00	21
TEACHER	0.75	0.92	0.83	26
Testing	1.00	1.00	1.00	37
Web Designing	1.00	1.00	1.00	24
accuracy			0.85	1145
macro avg	0.83	0.83	0.83	1145
weighted avg	0.84	0.85	0.84	1145



Дальнейшее развитие проекта.

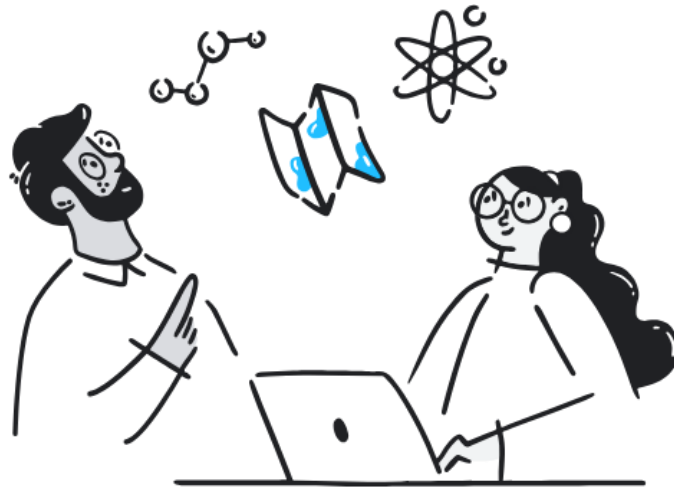
1. Перевод программы на русский язык.
2. Интерфейс программы.
3. Расширенный выбор профессий.





Выводы:

В своём проекте я выполнила сбор и анализ информации о профессиях и разработала программу с использованием искусственного интеллекта. Алгоритм был протестирован и дополнен. В работе программы не было выявлено ошибок.



Детский технопарк «Альтаир»



Список литературы.

1. Catboost. – URL: <https://catboost.ai/en/docs/> (дата обращения: 03.02.2024). – Режим доступа: свободный. – Текст: электронный.
2. Kaggle: Your Machine Learning and Data Science Community. – URL: <https://www.kaggle.com/datasets/> (дата обращения: 10.02.2024). – Режим доступа: для авториз. пользователей. – Текст: электронный.
3. Датасет 1: Kaggle: Resume Dataset. – URL: <https://www.kaggle.com/datasets/gauravduttakiit/resume-dataset/> (дата обращения: 03.02.2024). – Режим доступа: для авториз. пользователей. – Текст: электронный.
4. Датасет 2: Kaggle: Resume Dataset. – URL: <https://www.kaggle.com/datasets/snehaanbhawal/resume-dataset/> (дата обращения: 03.02.2024). – Режим доступа: для авториз. пользователей. – Текст: электронный.
5. Датасет 3: Kaggle: ResumeDataSet. – URL: <https://www.kaggle.com/datasets/dhainjeamita/resumedataset/> (дата обращения: 03.02.2024). – Режим доступа: для авториз. пользователей. – Текст: электронный.
6. Датасет 4: Kaggle: Updated Resume Dataset. – URL: <https://www.kaggle.com/datasets/jillanisofttech/updated-resume-dataset/> (дата обращения: 03.02.2024). – Режим доступа: для авториз. пользователей. – Текст: электронный.