Data Scientist CV Review Report

Thank you for downloading this customized CV analysis report. This report contains the outcome of a clustering analysis of your input skillset. We compared the skillset you provided to the skills required by hiring companies. This analysis is based on 21509 job requirements extracted from 2633 "Data Scientist" vacancy texts. This technique can be used to gain insight into how well a CV fits to the demand of employers. If you are looking for a job in data science this technique can be used to determine the most important skill gaps in your CV and to finetune your CV to employers demands. If you are a hiring manager or recruiter this technique can be used to complement the assessment of applicants CV's.

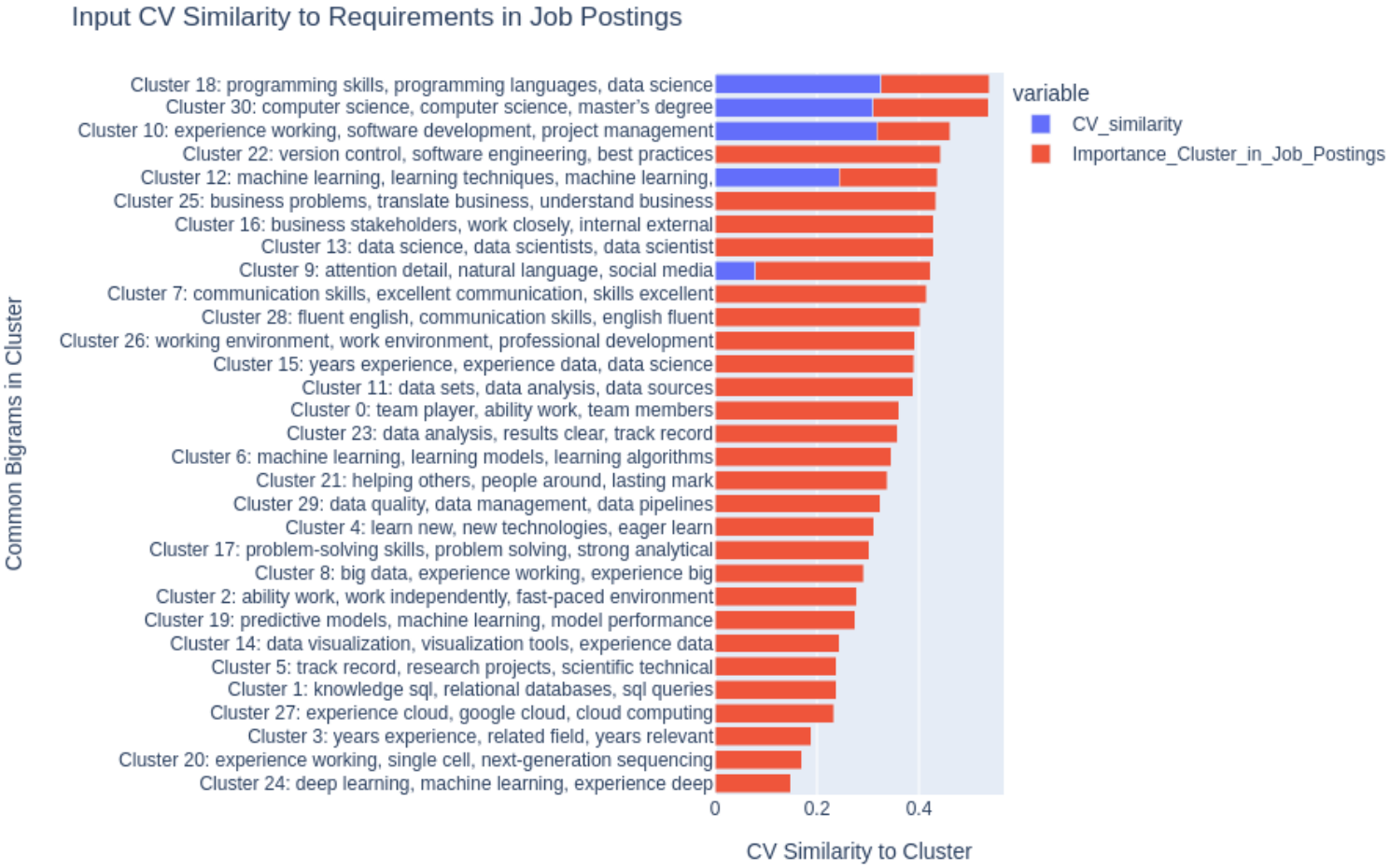
The methodology is based on a 2-step data pipeline where skills are processed using state of the art Natural Language Processing techniques (the Sentence-Transformers library), these processed skills are then compared to the most important skill clusters found in the dataset of 21509 job requirements. The methodology is developed as part of my thesis "Artificial Intelligence Driven Methods for the Analysis of Job Postings", accessible on https://github.com/koenbothmer/thesis.

Your Score: 0.52

Your total score is 0.52 This is an average score in comparison to a dataset of 65 Data Scientist CV's. Please note that allthough you score is average, more than 50% of Data Scientist CV's score better than yours. The mean score among these data scientist CV's is 0.49. The top 10% of these CV's scored 0.72. The top 25% of these CV's scored 0.67. The top 50% of these CV's scored 0.61.

Comparrison Plot

The clustering analysis of your input data is summarized in the plot below. The length of each red bar represents the importance which was determined by how often a cluster appears in vacancy texts.
The lenght of the blue bar represents how well your input skillset covers the red bar. Please note that it is not possbile to attain a score of 100%, a score of over 0.72% places in the top 10% of representation and can be considered excellent.



Analysis per input skill

Please find below the result of the analysis of each of your input skills.

Input Skill 1:
Your input skill "programming skills" was clustered in cluster 18 which contains skills regarding Cluster 18: programming skills, programming languages, data science. Your score for this skill is 0.6. this is above the average score among Data Scientist CV's which is 0.57

Input Skill 2:
Your input skill "computer science" was clustered in cluster 30 which contains skills regarding Cluster 30: computer science, computer science, master?s degree. Your score for this skill is 0.58.This score is quite low, the average score among Data Scientist CV's is 0.59 this may be due to a misclassification of our model but this could also indicate an opportunity to further clarify your CV.

Input Skill 3:
Your input skill "experience as a software developer" was clustered in cluster 10 which contains skills regarding Cluster 10: experience working, software development, project management. Your score for this skill is 0.69. this is above the average score among Data Scientist CV's which is 0.57

Input Skill 4:
Your input skill "using git and github" was clustered in cluster 9 which contains skills regarding Cluster 9: attention detail, natural language, social media. Your score for this skill is 0.19.This score is quite low, the average score among Data Scientist CV's is 0.21 this may be due to a misclassification of our model but this could also indicate an opportunity to further clarify your CV.

Input Skill 5:
Your input skill "machine learning techniques" was clustered in cluster 12 which contains skills regarding Cluster 12: machine learning, learning techniques, machine learning,. Your score for this skill is 0.56.This score is quite low, the average score among Data Scientist CV's is 0.65 this may be due to a misclassification of our model but this could also indicate an opportunity to further clarify your CV.

Your score compared to Data Scientist CV's

The table below shows an analysis of your scores are compared to an analysis of 65 Data Scientist CV's. This can be used to get an idea of where your CV needs improvement.
This might give hints on what learning activities to focus on, but also how well the skills you do possess are represented on your CV.

Skill Cluster	Your Score	Average	Top 10%	Top 25%	Top 50%
Cluster 0: team player, ability work, team members	0	0.61	0.74	0.74	0.73
Cluster 1: knowledge sql, relational databases, sql queries	0	0.55	0.7	0.68	0.64
Cluster 2: ability work, work independently, fast-paced environment	0	0.48	0.62	0.53	0.5
Cluster 3: years experience, related field, years relevant	0	0.61	null	null	0.68
Cluster 4: learn new, new technologies, eager learn	0	0.63	0.79	0.79	0.74
Cluster 5: track record, research projects, scientific technical	0	0.45	null	0.55	0.51
Cluster 6: machine learning, learning models, learning algorithms	0	0.53	0.68	0.65	0.61
Cluster 7: communication skills, excellent communication, skills excellent	0	0.73	0.85	0.83	0.78
Cluster 8: big data, experience working, experience big	0	0.57	0.76	0.71	0.66
Cluster 9: attention detail, natural language, social media	0.19	0.21	0.41	0.35	0.3
Cluster 10: experience working, software development, project management	0.69	0.57	0.71	0.67	0.63
Cluster 11: data sets, data analysis, data sources	0	0.51	0.69	0.64	0.59
Cluster 12: machine learning, learning techniques, machine learning,	0.56	0.65	0.77	0.75	0.72
Cluster 13: data science, data scientists, data scientist	0	0.62	0.73	0.71	0.68
Cluster 14: data visualization, visualization tools, experience data	0	0.6	0.78	0.75	0.7
Cluster 15: years experience, experience data, data science	0	0.7	0.82	0.79	0.76
Cluster 16: business stakeholders, work closely, internal external	0	0.57	0.71	0.68	0.64
Cluster 17: problem-solving skills, problem solving, strong analytical	0	0.7	null	null	0.74
Cluster 18: programming skills, programming languages, data science	0.6	0.57	0.72	0.69	0.65
Cluster 19: predictive models, machine learning, model performance	0	0.5	0.68	0.64	0.58
Cluster 20: experience working, single cell, next-generation sequencing	0	0.44	0.62	0.58	0.54
Cluster 21: helping others, people around, lasting mark	0	0.5	null	0.68	0.55
Cluster 22: version control, software engineering, best practices	0	0.45	0.63	0.59	0.54
Cluster 23: data analysis, results clear, track record	0	0.46	0.65	0.6	0.54
Cluster 24: deep learning, machine learning, experience deep	0	0.52	0.72	0.68	0.62
Cluster 25: business problems, translate business, understand business	0	0.51	0.68	0.64	0.59
Cluster 26: working environment, work environment, professional development	0	0.51	0.67	0.62	0.59
Cluster 27: experience cloud, google cloud, cloud computing	0	0.54	0.73	0.67	0.62
Cluster 28: fluent english, communication skills, english fluent	0	0.61	null	null	0.61
Cluster 29: data quality, data management, data pipelines	0	0.55	0.72	0.67	0.63
Cluster 30: computer science, computer science, master’s degree	0.58	0.59	0.69	0.68	0.66