A black and white logo

Description automatically generated with medium confidence

**AI Pricing research document**

OFS Platform

By Mohammed Al Harbi

Contents

[Introduction 3](#_Toc121113404)

[Goal 3](#_Toc121113405)

[Research questions 3](#_Toc121113406)

[1. Main research question: 3](#_Toc121113407)

[2. Sub-questions: 3](#_Toc121113408)

[Sub-questions findings 4](#_Toc121113409)

[1. What is the common category/vertical services among Oman governorates? 5](#_Toc121113410)

[**Method:** Stakeholder analysis 8](#_Toc121113411)

[2. What are the common determining factors of service prices among and within verticals? 9](#_Toc121113412)

[**Method:** Stakeholder analysis 10](#_Toc121113413)

[*3.* What is the best development approach for the artificial intelligence model? 10](#_Toc121113414)

[**Method:** Explore user requirements 11](#_Toc121113415)

[*4.* What are the different feasible algorithms for such an artificial intelligence model? 11](#_Toc121113416)

[**Method:** Available product analysis 12](#_Toc121113417)

[*5.* How can we evaluate the performance of the artificial intelligence model? 13](#_Toc121113418)

[**Method:** Available product analysis 13](#_Toc121113419)

[*6.* What are the scenarios in which the artificial intelligence model cannot be used? 13](#_Toc121113420)

[**Method:** Problem analysis 14](#_Toc121113421)

[Main-research question findings 14](#_Toc121113422)

[1. How can we utilize an artificial intelligence model efficiently to estimate the prices of freelancing services? 14](#_Toc121113423)

[References 15](#_Toc121113424)

# Introduction

Service pricing is always different and sometimes requires some experience to recognize whether it is overpriced, underpriced or the average market price. In the research, we will look at freelancing prevalent categories of services and their common pricing determining factors. It is very important to establish what those factors are in the hope of automating the process of evaluating freelancing service prices.

# Goal

The primary goal is achieved when some prices given their determining pricing factors can be predicted with use of artificial intelligence. The reason behind using an artificial intelligence model to solve this problem is because the process of evaluating prices requires an experienced person in the field of given category/vertical to be asked and it is possible that such people are hard to reach out to. Additionally, In OFS (Oman freelancing services platform) is it possible to produce freelancers data which consist of service description and price and this encourages the solution to be an artificial intelligence model.

# Research questions

# Main research question:

How can we utilize artificial intelligence technology efficiently to estimate the prices of freelancing services?

# Sub-questions:

What is the common category/vertical services among Oman governorates?

What are the common determining factors of service prices among and within verticals?

What is the best development approach for the artificial intelligence model?

What are the different feasible algorithms for such an artificial intelligence model?

How can we evaluate the performance of the artificial intelligence model?

What are the scenarios in which the artificial intelligence model cannot be used?

# Sub-questions findings

In this section, we will go over the answers to research sub-questions and eventually we can answer the main question based on those answers.

# What is the common category/vertical services among Oman governorates?

Having one fixed list is very difficult and one of the solutions to this is the possibility of adding freelancing services into the system by freelancers themselves. Below are some common verticals:

1. Delivering vertical:

This category of services is present in all Oman governorates and extremely important. Many truck drivers offer their item shipping services. Additionally, most Oman households use gas containers (also known as cylinders) to run cookstoves and use water storage tanks, which implies the importance of delivering gas containers and filling water tanks to what is known in Oman as “Tanker” which is a truck specialized and manufactured to deliver tons of water to households.

A picture containing lined, several

Description automatically generated

Gas containers

A picture containing sky, truck, outdoor, transport

Description automatically generated

Water tanker

1. Cleaning vertical:

These freelancing services are also very common in Oman, most often wives look for house cleaners during their work shift time and often those cleaners are available, but they do not know where to show themselves and the services they provide. Another type of service they offer is car cleaning, husbands commonly get back home tired and if the same person cleaning house also provides cleaning cars services, they often take both.

A picture containing floor, indoor, person

Description automatically generated

House cleaner

A picture containing calendar

Description automatically generated

Example of an advertisement about car washing freelancing service (which often not seen by the right audience)

1. Repairing and maintenance devices vertical:

Repairing mobile devices and laptops services can be offered by specialized companies and also by freelancers, often freelancers in this sector get very little attention mainly because of their poor advertisements regardless of their experience (Some situations customers could have the same performance of repairing, same duration but price between freelancer and a company can differ up to 100$ more) and often customers go for those companies because freelancers with such experience are not advertising themselves good enough.

1. Catering verticals:

In Oman typical occasions are celebrations and funerals. Catering services are heavily requested for those occasions. One example of freelancing service under catering is checking the quality of served food and ensuring the quantity is not over or under the needed to avoid wasting food.

1. Handyman verticals:

One of the most requested freelancing services are handyman services. Each handyman evaluates their prices and can differ from person to person by picking up bargains.

**DOT framework type**: Field

# **Method:** Stakeholder analysis

**Reason**: the type of the DOT framework is Field because this research question has been answered by the person living in Oman and experienced such services and the need for such a platform.

# What are the common determining factors of service prices among and within verticals?

To answer this question, we will take transportation and delivery verticals and understand the primary pricing determining factors, then we will compare them with different verticals such as construction.

The methodology implies whether common pricing factors lie among verticals, or every vertical has unique pricing factors for its services. Understanding where those factors lie is extremely vital as it plays a major role in the development approach of the artificial intelligence model.

Looking at the transportation and delivery vertical, one highly correlated pricing determining feature is whether natural disasters have happened recently or not. Typically, prior to those unfortunate events the related news is published then citizens get prepared. Afterwards, it is very likely that shipment and such services are unavailable, or inflation is likely to occur. Additionally, by investigating what differs water tankers delivery services from freelancer to another, the size of truck and how many gallons it possibly can carry that establishes the pricing disparity.

Reviewing the construction vertical, we found out that the freelancing services such architectural contractors and the pricing factors can be measured so for example the price of construction materials such as iron and cement, within the country play a huge role in pricing.

By evaluating the pricing factors of different verticals, we conclude that the factors within a vertical could have commonalities which we can take advantage of and utilize for improving an artificial intelligence model to estimate prices.

**DOT framework type**: Field

# **Method:** Stakeholder analysis

**Reason**: the type of the DOT framework is Field because this research question has been answered by the person living in Oman and experienced such services and the need for such a platform.

# What is the best development approach for the artificial intelligence model?

There are two common methodologies we can utilize to produce artificial intelligence models and picking the right approach plays a significant role in the overall efficiency of the model.

The first approach is known as the traditional approach where an artificial intelligence model gets exposed to all the data and attempts to spot patterns, in our case correlations to pricing and eventually we will end up with one model.

The great side of this approach is having to maintain one model is almost always easier than different approaches (same concept of monolith vs microservices apply.) On the other hand, you get less control over increasing the accuracy and have to rely solely on feedable data and algorithms.

The second approach is chaining artificial intelligence models (I personally like to call this method the micro-AI’s), this approach utilizes common rules such as generally the more choices you feed a model the less accurate it becomes, this is why data enrichment is extremely important as it can avoid this from happening in the traditional approach. However, given we acknowledge more choices equals less choices we can lift the problem one level higher where we train a model for each choice, given that choice in our case “verticals” and thus multiple artificial intelligence models will be created.

The great side of this approach is great control over breaking down problems into multiple smaller artificial intelligence models, each does a specific job. Applying accuracy performance techniques to each of those will drastically increase the overall accuracy and this combined approach combined with software allows for more accurate choice selection. On the other hand, multiple artificial intelligence models require most of the time maintenance and understanding the algorithms used, it is quite possible that background, collected data, techniques and algorithms used in one model differ totally to another which makes it more time consuming and hard to set up.

Since we have realized that each vertical can have their own freelancing services which are likely to have some commonalities, training an artificial intelligence on the data produced by freelancers within each vertical should result in maximizing the accuracy of the model. This is expected due to having to train an artificial intelligence model based on specific vertical data and thus no bias is taken into account by other verticals.

**DOT framework type**: Field

# **Method:** Explore user requirements

**Reason**: the type of the DOT framework is Field because the detailed view of reasoning behind the chosen methodology is clarified.

# What are the different feasible algorithms for such an artificial intelligence model?

Looking at the nature of the problem and the methodology we selected, it is quite clear that the problem is a regression problem where we attempt to predict a continuous value. We established that the methodology will be a chaining model but still, the parameters or features of each vertical does change from a vertical to another but there is always the text input from freelancers about their services and its price. Thus, the possibility of utilizing artificial intelligence in our case can occur using natural language processing techniques to predict prices, in other words, a natural language processing regression problem.

There are three primary types of algorithms for an artificial intelligence model that requires natural language processing. The first type is machine learning, one example of such a model is **MultinomialNB** offered by sci-kit learn library which specializes in machine learning models, usually those models are more transparent but not state-of-the-art techniques. The second type is deep-learning, one of the most effective deep learning architectures for natural language processing problems is **convolutional neural network (CNN)** which was heavily used on computer vision (for example, license plate recognition, object recognition and face recognition) and perform extremely well in natural language processing problems until 2017. The third type is transformers **BERT** (**Bidirectional Encoder Representations from Transformers)** This kind of models are considered the state-of-the-art in natural language processing problems and is published by google 2017-2018 those are quite complex models that trained on **huge datasets** and **trained for weeks/months** to obtain fantastic linguistic understanding and those are opened sources in hugging face library and available to be **tuned** to fit any problem; they are often big in size and by natural have bias because they trained with huge datasets, but can result in extremely high accuracy and performance.

**DOT framework type**: Library

# **Method:** Available product analysis

**Reason**: the type of the DOT framework is Library. We have researched existing research of different potential usable algorithms for price estimation.

# How can we evaluate the performance of the artificial intelligence model?

Since we have established that the problem we are solving is a regression problem, than we must use the regression evaluations metrics such as Mean Square Error (MSE)/Root Mean Square Error(RMSE) those can often tell on average what is the expected price in which the artificial intelligence model can overprice or underprice the estimation of a freelancing service. Additionally, Adjusted R Square can be utilized to explain the performance clearly to the audience and considers if the model is overfitted.

**DOT framework type**: Library

# **Method:** Available product analysis

**Reason**: the type of the DOT framework is Library. We have researched existing evaluation metrics for regression and their usage.

# What are the scenarios in which the artificial intelligence model cannot be used?

The artificial intelligence pricing estimation model will not be available during the data collection stage. We mentioned that having a fixed list of all verticals is very hard and freelancing services can be created by freelancers which allows us to utilize customer data to estimate prices per vertical. However, just like any artificial intelligence model having a good amount of data is important and so in the beginning of launching such a platform the AI services will most likely not be used until there are around 100-120 freelancer instances under the same vertical and the same freelancing service. This means that it is important to ensure customers can offer their freelancing services under existing verticals and freelancing services and so more data can be collected.

**DOT framework type**: Field

# **Method:** Problem analysis

**Reason**: the type of the DOT framework is Field since we investigate the problem of unavailable AI.

# Main-research question findings

*In this section, we will answer the main research question based on the answers to the sub-questions.*

# How can we utilize an artificial intelligence model efficiently to estimate the prices of freelancing services?

By establishing the fact that freelancing services fall under verticals or categories which most of the time have no common pricing factors, we have decided to take a chaining models approach where each category or vertical has its own artificial intelligence model. Freelancers can create new verticals, freelancing services or offer their own in an existing freelancing service. This way it is possible to use natural language processing to process the service description and estimate prices. As for algorithms and the methodology we took and the software architecture we have, aligns well with Transformers (BERT) models (we initially have a lack of data, thus utilizing powerful linguistic transformers should be the best option, aligns well with AWS SageMaker as those models can be big in disk space and take large RAM space).

# References

*Classifier Chain*. (n.d.). Scikit-learn. https://scikit-learn.org/stable/auto\_examples/multioutput/plot\_classifier\_chain\_yeast.html

*Convolutional Neural Networks (CNN) with Deep Learning*. (2020, January 28). HappiestMinds. https://www.happiestminds.com/insights/convolutional-neural-networks-cnns/

*Hugging Face – The AI community building the future.* (n.d.). https://huggingface.co/

*Working With Text Data*. (n.d.). Scikit-learn. https://scikit-learn.org/stable/tutorial/text\_analytics/working\_with\_text\_data.html

Wu, S. (2021, December 14). *3 Best metrics to evaluate Regression Model? - Towards Data Science*. Medium. https://towardsdatascience.com/what-are-the-best-metrics-to-evaluate-your-regression-model-418ca481755b