Designing of an AI system for predicting the successful recovery of SpaceX's Falcon 9 launch vehicle

1Sandesh Yadav, 2Yash Manmode, 3Mansi Patil
1Student, 2Student, 3Student
G. H. Raisoni Institute of Engineering and Technology, Nagpur

Abstract - **This paper demonstrates the usage of various ML algorithms including neural network (AI) for the Falcon 9 Launch Vehicle's first-stage Recovery Prediction System. We predicted the success of the Falcon 9 first-stage landing. On its website, SpaceX promotes Falcon 9 rocket launches for 62 million dollars; other suppliers charge upwards of 165 million dollars for each launch. A large portion of the savings is due to SpaceX's ability to reuse the first stage. So, if we can determine whether the first stage will land, we can determine how much a launch will cost. If a different business wants to compete with SpaceX for a rocket launch, it may use the information provided here. SpaceX's overarching objective is to make Falcon 9's first stage reusable by returning it to the landing zone following separation. The accuracy of Falcon 9 launch vehicle recovery is predicted using a variety of AI models, including Logistic Regression, SVM, Decision Tree, Random Forest, KNN, and ANN.**

***keywords* - Launch Recovery System, AI, SpaceX, Falcon 9**

I. INTRODUCTION

The development of technology never stops. Designing a product with the help of modern technology that will improve the lives of others is a significant contribution to the neighbourhood and environment. Our approach uses artificial intelligence to predict Falcon Launch Vehicle recovery. The Falcon9 is a medium-lift launch vehicle that uses partially two stages to reach orbit and was developed and produced in the US by SpaceX. The SpaceX Falcon 9 is a reusable, two-stage rocket that can carry passengers and payloads reliably and safely into Earth orbit and beyond. The first orbital-class reusable rocket in the world is called Falcon 9. Reusability enables SpaceX to relaunch the rocket's most expensive components, lowering the cost of space access.

According to SpaceX's website, Falcon 9 rocket flights cost 62 million dollars, compared to upwards of 165 million dollars for each launch from other companies. A large portion of the savings comes from the ability to reuse the first stage. Therefore, we can calculate the price of a launch if we can estimate if the first stage will land. If another business wants to compete with SpaceX for a rocket launch, it can utilize this information. The only business in the world that can reliably reuse rocket boosters of the orbital class is SpaceX. With the help of its fleet of reusable first-stage Falcon 9 rockets, SpaceX has been able to lower the cost of space travel. The company has launched 148 missions and 110 times by landing on autonomous seagoing droneships and landing pads, orbital-class rocket boosters have been retrieved. It has reused Falcon 9 first-stage boosters 87 times out of the rockets that have been found.

Objectives:

- Determining the cost of SpaceX's Falcon 9 launch by predicting whether the Falcon 9 launch vehicle will recover its first stage.
- It will eventually reduce fuel consumption while optimally utilizing the precious non-renewable resources provided by nature
- Even though the design is more expensive it can be reused frequently, and in the long run, it's more cost-effective.
- The following reason, it reduces the stress on expenses. In due course, it will lead to a depletion in E-waste and space debris.

II. LITERATURE SURVEY

"Falcon 9 Launch Vehicle" The SpaceX Falcon 9 is a reusable, two-stage rocket that can carry passengers and payloads safely and reliably into Earth orbit and beyond. The first orbital-class reusable rocket in the world is called Falcon 9. Reusability enables SpaceX to reuse the rocket's most expensive components, lowering the cost of space access.

"Fuel Conservation for Launch Vehicles: Falcon Heavy Case Study" Primoz Jozic, Aleksander Zidansek, Robert Repnic . Published on February 20. This paper proposes that Recent growth in space exploration has had a negative impact on the environment, especially when rockets are launched since a lot of fuel is used and exhaust gases are expelled into the atmosphere. In this case study, one of the most promising rockets for the reasonably priced carrying of large payloads into orbit and beyond is the SpaceX Falcon Heavy reusable heavy-lift launch vehicle. For the instance of its maiden launch on February 6, 2018, we assessed several options for optimizing fuel use and minimizing environmental effects. [1]

"Reusable launch rocket Challenges And Control Problems: Falcon 9 launch vehicle" Alexander Nebylov, Vladimir Nebylov published on December 16. In this paper the possible directions of development of space launch technology, including space launch to a suborbital trajectory, in order to reduce the specific cost of launch at the expense of the majority of reusable carrier components, are analysed. Opportunities for providing reusability for horizontal and vertical launches are compared. The experience of the soft landing of the first stage of the Falcon 9-R rocket of US firm SpaceX is taken into account. The requirements of the air-breathing engine, which could provide an economical horizontal launch, are considered. The requirements for the engine could be reduced for suborbital launch, and in this simplified case, they could be fulfilled well already at the present stage of development of aerospace technologies. [2]

"Launch Vehicle Recovery and Reuse: History of reusable launch vehicles ". Over a number of years, SpaceX has been working on technologies that would enable complete and quick reusability of space launch vehicles. The long-term goals of the project include returning the first stage of a launch vehicle to the launch site in a matter of minutes and returning the second stage to the launch pad after orbital realignment with the launch site and atmospheric re-entry, which might take up to 24 hours. Long-term, SpaceX wants both of its orbital launch vehicle's stages to be capable of being reused a few hours after landing. For Falcon 9's first stage, the reusable launch system technology was created and first employed. The booster turns around after stage separation, an optional boost back burn is performed to change its direction, a re-entry burn is performed, controlling direction to reach the landing site, and a landing burn is performed to impact the final low-altitude deceleration and touchdown. Because the vehicle is moving at orbital velocity, extending reusable flight hardware to the second stages is a more difficult engineering challenge that SpaceX planned to tackle (at least starting in 2014). Initial ideas to make Falcon 9's second stage reusable were dropped. By 2021, SpaceX plans to replace all of its current launch vehicles and spacecraft used for satellite delivery and passenger transportation—Falcon9—with the Starship system, a completely reusable two-stage launch vehicle. [3]

"Characterization of Falcon 9 launch vehicle noise from far-field measurements" Logan Mathews, Kent Lee, Grant Hart Published on July 2021. In this study, the Falcon 9, a cutting-edge launch vehicle with a rapid operational pace, is examined for its source-related noise characteristics. Long-standing research has focused on the empirical prediction of the noise properties of launched rockets; nevertheless, there are only a small number of comparisons with high-fidelity, far-field data, and historical discrepancies still exist. The following factors are taken into account: peak frequency, maximum overall sound pressure level (OASPL), overall directivity, and overall sound power. Although the historical literature indicates that the noise directivity of the Falcon 9 vehicle is between two different ranges, convective Mach number theories provide a good representation of the reported peak directivity angle. The radiation efficiency obtained from comparing mechanical and acoustic power is in agreement with the existing literature. Even at distances of many kilometres, two different techniques of forecasting the maximum OASPL provide findings that are accurate to within 2 dB. For the obtained spectral peak frequency, several scaling parameters are determined to relate these findings to earlier observations. The effect of terrain shielding on levels and spectra is evaluated in the last section. These Falcon 9 vehicles discovered source characteristics offer a link to earlier launch vehicle acoustics research, making it easier to find practical models and approaches for comprehending rocket noise. [4]

III. METHODOLOGY

The Methodology for this artificial intelligence system includes the following steps:

1. **Data Collection:** In this phase, we collect the data from SpaceX REST API which is an open-source API specially maintained by SpaceX officials to display launch data publically. This API will give us data about launches, including information about the booster version used, payload delivered, launch specifications, landing specifications, and landing outcomes.

2. **Storing:** After data is gathered, the data is saved in the MySQL database, an open-source relational database management system (RDBMS) built on the Structured Query Language and supported by Oracle (SQL). It works with almost any operating system, including Linux, UNIX, and Windows.

3. **Data Cleaning:** This stage involves cleaning up the data by locating faults or corruptions, fixing them or erasing them, adding or removing null values, or manually processing the data as necessary to avoid repeating the same mistakes.

4. **Data Preprocessing:** Python and Microsoft Excel were both used for the data preparation. Preprocessing the data is crucial since the raw data that has been obtained cannot be analyzed by a machine learning algorithm.

5. **Exploratory Data Analysis (EDA):** Data wrangling, or EDA, is done before moving on to predictive analytics to extract valuable information from the data that a human eye could miss. MySQL is used to do data wrangling and EDA, and Python and Tableau are used to build visualizations.

6. **Normalization**: Once a descriptive analysis has been completed, a predictive analysis should be undertaken. We must normalize the data to ensure that all of the variables fall within the same range before building and training the models. If not, the model can start to bias.

7. **Model Creation**: After data has been normalized or standardized, various Deep learning and machine learning-based algorithms are applied to the data. Several models are employed, including K Nearest Neighbors (KNN), Logistic Regression, Decision Tree, Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks.

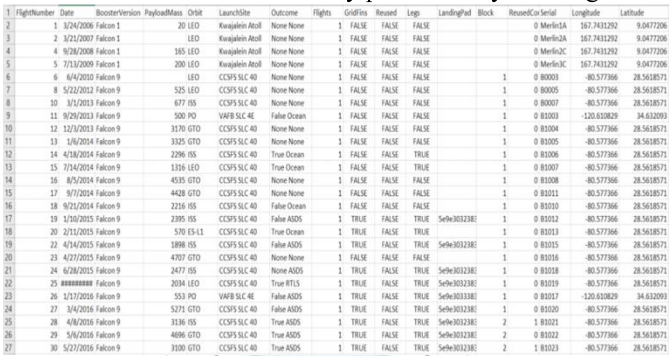
8. **Hyperparameter Tuning**: The parameters of the models must be hyper-tuned once they are developed to provide the greatest accuracy.

9. **Model Selection**: After each model has undergone hyper-tuning, it is time to choose the one that provides the highest accuracy and best fits the situation at hand.

The flowchart of gathering data from SpaceX API is given below,

***Fig 1**: Flowchart of gathering data from SpaceX API.*

The project workflow for Falcon 9 Launch Vehicle recovery prediction system is given below,



1	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	ReusedCo	Serial	Longitude	Latitude
2	1	3/24/2006	Falcon 1	20	LEO	Kwajalein Atoll	None None	1	FALSE	FALSE	FALSE			0	Merlin1A	167.7431292	9.0477206
3	2	3/21/2007	Falcon 1		LEO	Kwajalein Atoll	None None	1	FALSE	FALSE	FALSE			0	Merlin2A	167.7431292	9.0477206
4	4	9/28/2008	Falcon 1	165	LEO	Kwajalein Atoll	None None	1	FALSE	FALSE	FALSE			0	Merlin2C	167.7431292	9.0477206
5	5	7/13/2009	Falcon 1	200	LEO	Kwajalein Atoll	None None	1	FALSE	FALSE	FALSE			0	Merlin3C	167.7431292	9.0477206
6	6	6/4/2010	Falcon 9		LEO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B0003	-80.577366	28.5618571
7	8	5/22/2012	Falcon 9	525	LEO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B0005	-80.577366	28.5618571
8	10	3/1/2013	Falcon 9	677	ISS	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B0007	-80.577366	28.5618571
9	11	9/29/2013	Falcon 9	500	PO	VAFB SLC 4E	False Ocean	1	FALSE	FALSE	FALSE		1	0	B1003	-120.610829	34.632093
10	12	12/3/2013	Falcon 9	3170	GTO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B1004	-80.577366	28.5618571
11	13	1/6/2014	Falcon 9	3325	GTO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B1005	-80.577366	28.5618571
12	14	4/18/2014	Falcon 9	2296	ISS	CCSFS SLC 40	True Ocean	1	FALSE	FALSE	TRUE		1	0	B1006	-80.577366	28.5618571
13	15	7/14/2014	Falcon 9	1316	LEO	CCSFS SLC 40	True Ocean	1	FALSE	FALSE	TRUE		1	0	B1007	-80.577366	28.5618571
14	16	8/5/2014	Falcon 9	4535	GTO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B1008	-80.577366	28.5618571
15	17	9/7/2014	Falcon 9	4428	GTO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B1011	-80.577366	28.5618571
16	18	9/21/2014	Falcon 9	2216	ISS	CCSFS SLC 40	False Ocean	1	FALSE	FALSE	FALSE		1	0	B1010	-80.577366	28.5618571
17	19	1/10/2015	Falcon 9	2395	ISS	CCSFS SLC 40	False ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	1	0	B1012	-80.577366	28.5618571
18	20	2/11/2015	Falcon 9	570	ES-L1	CCSFS SLC 40	True Ocean	1	TRUE	FALSE	TRUE		1	0	B1013	-80.577366	28.5618571
19	22	4/14/2015	Falcon 9	1898	ISS	CCSFS SLC 40	False ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	1	0	B1015	-80.577366	28.5618571
20	23	4/27/2015	Falcon 9	4707	GTO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B1016	-80.577366	28.5618571
21	24	6/28/2015	Falcon 9	2477	ISS	CCSFS SLC 40	None ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	1	0	B1018	-80.577366	28.5618571
22	25	#########	Falcon 9	2034	LEO	CCSFS SLC 40	True RTLS	1	TRUE	FALSE	TRUE	5e9e3032383	1	0	B1019	-80.577366	28.5618571
23	26	1/17/2016	Falcon 9	553	PO	VAFB SLC 4E	False ASDS	1	TRUE	FALSE	TRUE	5e9e3033383	1	0	B1017	-120.610829	34.632093
24	27	3/4/2016	Falcon 9	5271	GTO	CCSFS SLC 40	False ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	1	0	B1020	-80.577366	28.5618571
25	28	4/8/2016	Falcon 9	3136	ISS	CCSFS SLC 40	True ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	2	1	B1021	-80.577366	28.5618571
26	29	5/6/2016	Falcon 9	4696	GTO	CCSFS SLC 40	True ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	2	0	B1022	-80.577366	28.5618571
27	30	5/27/2016	Falcon 9	3100	GTO	CCSFS SLC 40	True ASDS	1	TRUE	FALSE	TRUE	5e9e3032383	2	1	B1023	-80.577366	28.5618571

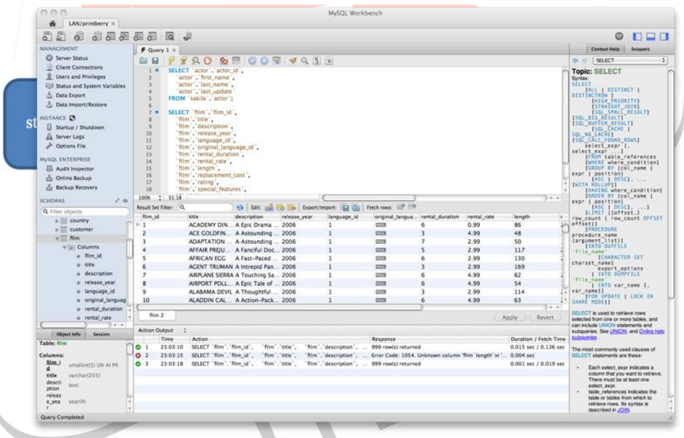


***Fig 2**: Project workflow for Falcon 9 Launch Vehicle recovery prediction system*

And the Use-Case diagram of the Prediction system is also given below,

***Fig 3**: Use-Case diagram of the AI Prediction system*

IV. IMPLEMENTATION

The following steps were performed to make the AI predictive system as listed in the methodology above.

1. **Data Collection**: The data is collected from the SpaceX REST API is an open-source repository that contains all the data for SpaceX launch, rocket, core, capsule, Starlink, launchpad, and landing pad data. The data contains variables like Booster version, payload mass, launch orbit, longitude, latitude, etc. Since SpaceX's launch recovery system technology is still relatively new to the world, the data available for it is quite less. The data set contains 17 variables/features and 98 records.

***Fig 4**: Data collected in .csv file for building model*

2. Data Storage: The collected data was stored using MS Excel and MySQL database, which is a very famous open-source relational database management system, which is mainly used for storing data and querying. Since the data available and collected through SpaceX REST API is mostly structured, storing it in a relational database was the best option MySQL is best for the job.

***Fig 5**: Data Stored in MySQL Database*

3. **Data Cleaning:** The data collected from SpaceX REST API is raw data and thus mostly unclean and unprocessed. It contains a lot of null values, and errors and is a format that may not be processed by machine learning and deep learning models. Thus it is important to clean this data first.

To deal with missing values, first, we need to find how many missing values each column/features contain. To do this the isnull() and sum() functions are used. After getting the count of how many numerical values we have missing, we usually fill them with the mean of all the values of the column. This done is using the fillna() function. The mean of the column is passed as an argument in the function.

4. **Data Wrangling:** Data wrangling or simply data pre-processing is the process of converting data from raw form to processable form i.e. to a form that is processable by the ML and DL algorithms. Our dataset contains many categorical and mixed features like launch orbit, booster version, etc. These variables are needed to be one-hot encoded or converted into binary form. But to do all this, first, we need to find out what and how many columns contain categorical features. This can be done by dtypes() method.
Since the launch recovery system technology is relatively new, the tests and data for this are considerably low for analysis purposes. Therefore, our priority was to add dummy data to our present dataset in order to train the model with an appropriate amount of data. As deep learning models like ANN or artificial neural networks require high amounts of data for training and optimization purposes, this was a crucial step. Adding dummy data to our original dataset increased the amount of data for training and analysis dramatically.
5. **Exploratory Data Analysis (EDA):** Exploratory data analysis or EDA is done to extract necessary information from the data that can be used to answer various questions and hypotheses that we assume during the project. Visualizations are the main part of EDA, since graphs and visualizations and easy to understand and depict, they are used to understand information that can be missed by the human eye. The visualization libraries used in this project are Seaborn and Matplotlib. For EDA, we used the following graphs and charts for data exploration:
   (1) Scatter graphs:
      (i) Flight Number Vs. Payload Mass
      (ii) Flight Number Vs. Launch Sites
      (iii) Orbit Vs. Flight Number
      (iv) Payload Vs. Flight Number
      (v) Payload Vs. Orbit Type
      (vi) Orbit Vs. Payload
   (2) Bar Graphs:
      (i) Mean Vs. Orbit
   (3) Line Graphs:
      (i) Success Rate Vs. Year

   After the complete exploratory data analysis using the given graphs and charts, we came to the following conclusions:
   - The success rate of a launch site is directly proportional to the number of flights at that site.
   - For the VAFB-SLC launch site, there are no rockets launched for heavy payload masses (greater than 10000 kg).
   - Orbits ES-L1, GEO, HEO, and SSO have the highest success rate.
   - In the LEO orbit, the success appears to be related to the number of flights; on the other hand, there seems to be no relationship between flight numbers when in the GTO orbit.
   - With heavy payloads, the successful landing or positive landing rate is more for PO, LEO, and ISS.
   - However, for GTO, we cannot distinguish this well as both positive landing rates and negative landings are there.
   - The success rate from 2013 to 2020 kept increasing.
   - The success rate of low-weighted payloads is higher than high-weighted payloads.
6. **Standardization of data and Splitting into Training and Testing data:** After all the pre-processing and cleaning, our next step is Standardize the data i.e. to convert all features into a similar scale. This is a very crucial step and if skipped, it may lead to our models being biased. There are many ways and functions available to standardize or normalize the data. We used StandardScaler() method present in the sklearn library.
It's crucial to divide the data into training and testing after scaling. The testing data is used to test the model once it has been developed, whereas the training data is used to train the ML/DL model. To split the data into a 4:1 ratio, we used the train_test_split() method.
7. **Model Creation:** Once all the standardization and splitting are done, machine learning and deep learning algorithms were applied to the dataset for predictive analysis. The libraries used for this purpose are sklearn, TensorFlow and Keras. The ML and DL algorithms that were used in the projects are:-
   I. Logistic Regression
   II. Support Vector Machines (SVM)
   III. Decision Tree
   IV. Random Forest
   V. K Nearest Neighbors (KNN)
   VI. Artificial Neural Networks (ANN)
8. **Hyperparameter Tuning:** Hypermeter tuning, as the name suggests, refers to tuning the parameters of the ML and DL algorithms concerning the given data so that we can get the maximum possible accuracy. For automated hyper tuning of the algorithms, we used the GridSearchCV() function. GridSearchCV() is a special type of function that tests the algorithms for all the combinations of parameters provided as an argument into the function, and also applies a cross-validation algorithm on the dataset. The parameters that are to be tested are provided in the form of the dictionary to the function.

V. RESULTS

In this Project, we created a project to project to predict the price of SpaceX's Falcon 9 launch vehicle by predicting whether the Falcon 9 launch vehicle can recover its first stage. After all the cleaning, pre-processing, standardizing, splitting the data, and hyper tuning all the algorithms, the Artificial Neural Networks turned out to be the best model, which gave an accuracy of a 100% on the testing data.

Fig 6*: Summary of Artificial Neural Network (ANN) Model*

After the model are built, we tested the accuracy of various models,

Fig 7*: Accuracy of all models*

Fig 8*: Accuracy visualization of all models*

In this work, the ANN is proved to be the best model for the prediction of the Falcon 9 launch vehicle which gives 100% accuracy that it is recovered in its first stage successfully. And the price will estimate reduce up to 62 million dollars

```
Model: "sequential"
_________________________________________________________________
Layer (type)                Output Shape              Param #
=================================================================
dense (Dense)               (None, 30)                2520
_________________________________________________________________
dense_1 (Dense)             (None, 25)                775
_________________________________________________________________
dense_2 (Dense)             (None, 20)                520
_________________________________________________________________
dense_3 (Dense)             (None, 15)                315
_________________________________________________________________
dense_4 (Dense)             (None, 10)                160
_________________________________________________________________
dense_5 (Dense)             (None, 5)                 55
_________________________________________________________________
dense_6 (Dense)             (None, 1)                 6
=================================================================
Total params: 4,351
Trainable params: 4,351
Non-trainable params: 0
_________________________________________________________________
```

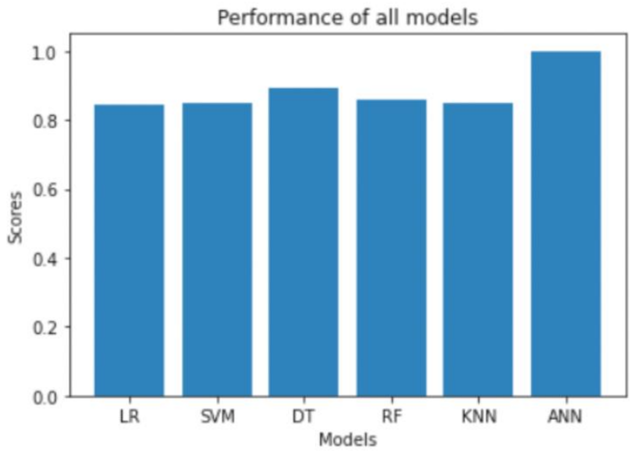
from 165 million dollars. This information can be beneficial to an alternate companies wants to bid against SpaceX for the rocket Launch.

VI. FUTURE SCOPE

The organizations working on reusable technology are creating several opportunities in the near future within the context of

	Model	Score
0	Logistic Regression	0.846429
1	Support Vector Machine	0.848214
2	Decision Tree	0.891071
3	Random Forest	0.860714
4	K Nearest Neighbor	0.848214
5	Artificial Neural Networks	1.000000

space exploration. This technology can possibly increase the frequency of launches, thereby increasing the scope of operations.



Space tourism is one of the upcoming sectors in the space industry — human space travel for recreational purposes and has gained some traction in the past few years. Various aerospace companies, other than SpaceX such as Blue Origin, and Virgin Galactic, have started developing space tourism systems. There are different types of space tourism, namely, lunar, orbital, and suborbital space tourism. Companies such as Blue Origin and SpaceX are actively working on reusable launch systems to support space tourism.

Reusable launch systems are also known as reusable launch vehicles as they cater to the recovery of some of the major components of the launch system. The reusability factor of these systems has generated a great deal of interest and has further fuelled technological development in this area. These systems hold a huge potential market in the future as they can drastically reduce the investment associated with space launches and exploration.

The entry of Space X into the reusable rocket segments has ignited a technical innovation in the market and has pushed the competitors to invest in launch industry technological innovations. These advancements have led to significant changes in the global launch industry, both in terms of supply and demand along with technological revolution caused by existing companies and new entrants in the market.

VII. REFERENCES

1. Primoz Jozic, Aleksander Zidansek, Robert Repnic, Faculty of Natural Sciences and Mathematics, Physics Department, University of Maribor, Koroš ka cesta 160, SI-2000 Maribor, Slovenia.
2. Alexander Nebylov, Vladimir Nebylov, State University of Aerospace Instrumentation, 67 Bolshaya Morskaya, Saint-Petersburg, 190000 Russia.
3. Simberg, Rand (February 8, 2012). "Elon Musk on SpaceX's Reusable Rocket Plans" from the original on October 6, 2014. Retrieved February 7, 2012.
4. Logan T. Methews, Kent L. Gee, Grant W. Hart, Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA.