■ Lung Cancer Risk Prediction - README

■ Lung Cancer Risk Prediction

This project predicts the risk of lung cancer based on user data such as age, gender, smoking habits, and other health indicators.
It applies machine learning techniques to identify individuals at high risk.

■ Features
- Data preprocessing and cleaning
- Exploratory Data Analysis (EDA)
- Model building and evaluation (Logistic Regression / Random Forest)
- Visualization of results
- Predictive web app using Streamlit (optional)

■ Dataset
The dataset includes features such as:
- Age
- Gender
- Smoking
- Yellow fingers
- Anxiety
- Peer pressure
- Chronic disease
- Fatigue
- Allergy
- Wheezing
- Alcohol consumption
- Coughing
- Shortness of breath
- Swallowing difficulty
- Chest pain
- Lung cancer (target variable)

■ Machine Learning Model
Used algorithms:
- Logistic Regression
- Random Forest Classifier

Performance metrics:
- Accuracy
- Precision
- Recall
- F1-Score

■■ Installation

Clone this repository:
```
git clone https://github.com//lung-cancer-risk-prediction.git
cd lung-cancer-risk-prediction
```

Install dependencies:
```
pip install -r requirements.txt
```

Run the model:
```
python src/train_model.py
```

(Optional) Launch the app:
```
streamlit run app.py
```

■ Results
Model achieved an accuracy of around **X%** on test data.

■■■ Author
****
B.Tech CSE (Data Science)
Madanapalle Institute of Technology & Science

■ License