Kyoto	Kyoto	Kyoto Kyoto	Kyoto
Lagos	Lagos	Lagos Lagos	Lagos
Lagos	Lagos	Lagos	Lagos
Lagos	Lagos	Lagos	Lagos
Lagos	Lagos	Lagos Kyoto Kyoto	Kyoto
Lagos	Lagos	Lagos	Lagos
Lagos	Lagos	Lagos	Lagos
Lagos	Lagos	Lagos	Lagos

Disease	Symptoms	Treatment
Asthma	coughing, shortness of breath, wheezing	
Pneumonia	fatigue, wheezing, fever	
Bronchitis	Low Energy, Wheezing, coughing	
Asbestosis		
Tuberculosis		
Aspergillosis		
pneumothorax		

IDEA of what LDA will give: Rows represent the text, columns represents the topics that LDA thinks are important. Recall that LDA uses TF-IDF or BoW

	coughing	Low energy	wheezing	.....
Text 1	4.5	1.2	0.6	6.3
Text 2	0.3	4.7	1.1	
Text 3	2.5	5	6	
....				

	Multi-label	**Single Label**	
	Top N Labels (N=3)	Max Label	
1	Bronchitis, Asthma, Pneumonia	Bronchitis	1 blue
2	Asthma, Bronchitis, Pneumonia	Asthma	2 red
3	Asthma, Pneumonia, Bronchitis	Asthma	2 red

We want to sum up scores if we see that the the topics LDA gives are in the list of symptoms or treatments. Then use that weight score as a likelihood value that the text corresponds to that disease.

	Asthma	Pneumonia	Bronchitis	....	Top 3:	Bronchitis
Text 1	5.1	0.6	6.3			Asthma
						Pneumonia
Text 2	Asthma	Pneumonia	Bronchitis			
	#	#	#			Asthma
Text 3	Asthma	Pneumonia	Bronchitis			Asthma
	#	#	#			

So once we get the scores, we are able to choose the top N diseases in ranking order of scores in order to determine what disease is most likely and then use that as the labels to feed into the Classification Model. Note, Classification Models need a label in order to work otherwise we have to work in the realm of unsupervised learning. For simplicity, I think we shoudl use this approach of topic modelling with LDA to formulate the labels.

Conversation Dataset		Conversation	Diagnosis (Top N)	OR	Diagnosis (MAX Label)
	0	gjsgoesgoksh	Bronchitis, Asthma, Pneumonia		Bronchitis
	1	sdgdsgshdhwsh	Asthma, Bronchitis, Pneumonia		Asthma
	2	whwhwrhwrh	Asthma, Pneumonia, Bronchitis		Asthma
	3	wrhwhwrhwr			
....					
	212	wrhwhwrhwr			

Classification Model

Training set	80 % of the conversations	Label goes into the ML model
Testing Set	20% of conversations	Label is not given into the ML models. It is isolated because when we use the ML model on the testing set it will predict what disease it will be after learning off the training set. Then we compare what the model predicts to the labels we isolated and do the folllowing metrics:

Ways to assess performance of ML model

1. Confusion Matrix: Accuracy, TP, TN, FP, FN
2. AUC: Area under the Curve
3. AIC: information criterion

Disease	Symptoms	Treatment	Summary
Acute Respiratory Distress Syndrome	Shortness of breath, Low blood pressure, Confusion, Extreme Tiredness, Labored Breathing, Rapid Breathing	inhaler, oxygen, mechanical ventilation	Acute respiratory distress syndrome (ARDS) occurs when fluid builds up in the tiny, elastic air sacs (alveoli) in your lungs. The fluid keeps your lungs from filling with enough air, which means less oxygen reaches your bloodstream. This deprives your organs of the oxygen they need to function.
Asbestosis	Fingertips and toes that appear wider and rounder than usual, Shorness of breath, Chest pain, Chest Tightness, Dry Cough, Crackling sounds	stay away from cold places, oxygen, therapy, surgery, lung function tests	Asbestosis (as-bes-TOE-sis) is a chronic lung disease caused by inhaling asbestos fibers. Prolonged exposure to these fibers can cause lung tissue scarring and shortness of breath. Asbestosis symptoms can range from mild to severe, and usually don't appear until many years after initial exposure.
Aspergillosis	Fever, Produce mucus, Runny nose, Stuffy nose	itraconazole, surgery, antifungal medications, embolization, observation	Aspergillosis is an infection caused by a type of mold (fungus). The illnesses resulting from aspergillosis infection usually affect the respiratory system, but their signs and severity vary greatly.
Asthma	Coughing, shortness of breath, Wheezing, Chest Tightness, Chest Pain, Trouble Sleeping	Omalizumab, Mepolizumab, inhaler	Asthma is a condition in which your airways narrow and swell and may produce extra mucus. This can make breathing difficult and trigger coughing, a whistling sound (wheezing) when you breathe out and shortness of breath.
Bronchiolitis	Stuffy nose, Runny nose, Cough, Fever	Aerosolized ribavirin, saline nose drops, nebulized albuterol, corticosteroid, chest physiotherapy	Bronchiolitis is a common lung infection in young children and infants. It causes swelling and irritation and a buildup of mucus in the small airways of the lung. These small airways are called bronchioles. Bronchiolitis is almost always caused by a virus.
Bronchitis	Coughing, Fatigue, Produce mucus, Fever, Shortness of breath, chills	antibiotic, cough medicine, inhaler	Bronchitis is an inflammation of the lining of your bronchial tubes, which carry air to and from your lungs. People who have bronchitis often cough up thickened mucus, which can be discolored. Bronchitis may be either acute or chronic.
Chronic Bronchitis	Chest pain, Coughing, Fatigue, Produce mucus, Fever	steroids to reduce inflammation, antibiotic, cough medicine, inhaler	Bronchitis is an inflammation of the lining of your bronchial tubes, which carry air to and from your lungs. People who have bronchitis often cough up thickened mucus, which can be discolored. Bronchitis may be either acute or chronic.
Chronic obstructive pulmonary disease	Swelling in ankles, Shortness of breath, Wheezing, chest pain, low energy, weight loss, chest tightness, chronic cough, infections, low energy, weight loss,	Pulmonary rehabilitation, Bronchodilators, Albuterol , Ipratropium , Levalbuterol, inhaled steroids, inhalers, antibiotics, Theophylline, oxygen therapy	lung disease that causes obstructed airflow from the lungs. Symptoms include breathing difficulty, cough, mucus (sputum) production and wheezing. It's typically caused by long-term exposure to irritating gases or particulate matter, most often from cigarette smoke. People with COPD are
Influenza	Headache, Sore throat, Stuffy nose, Runny nose, low energy, dry cough, tiredness, weakness, eye pain, vomiting, diarrhea	Oseltamivir, rest, fluids, Relenza, Tamiflu, Xofluza	Flu, also called influenza, is an infection of the nose, throat and lungs, which are part of the respiratory system. The flu is caused by a virus. Influenza is commonly called the flu, but it's different from the stomach "flu" viruses that cause diarrhea and vomiting.
Pneumonia	Fatigue, Wheezing, Fever, Sweating, Nausea, chest pain, confusion, cough, chills, phlegm, shortness of breath, vomiting, diarrhea	antibiotic, asprin, cough medicine, ibuprofen, Advil, Motrin, Tylenol, acetaminophen	Pneumonia is an infection that inflames the air sacs in one or both lungs. The air sacs may fill with fluid or pus (purulent material), causing cough with phlegm or pus, fever, chills, and difficulty breathing. A variety of organisms, including bacteria, viruses and fungi, can cause pneumonia.
Pneumothorax	Chest pain, Shortness of breath	isotonic sodium chloride solution, needle aspiration, chest tube, oxygen therapy, observation	A pneumothorax (noo-moe-THOR-aks) is a collapsed lung. A pneumothorax occurs when air leaks into the space between your lung and chest wall. This air pushes on the outside of your lung and makes it collapse. A pneumothorax can be a complete lung collapse or a collapse of only a portion of the lung.
Respiratory syncytial virus	Headache, Runny nose, Dry cough, congestion, fever, sore throat, sneezing	Intravenous fluids, acetaminophen, Tylenol, nasal saline drops, suctioning, antibiotics, oxygen, ventilation, IV, inhaler	Respiratory syncytial virus (RSV) causes infections of the lungs and respiratory tract. It's so common that most children have been infected with the virus by age 2. Respiratory syncytial (sin-SISH-ul) virus can also infect adults.
Tuberculosis	Fever, Low energy, Coughing, chest pain, night sweats, fatigue, tiredness, weight loss, chills	rifampin, pyrazinamide, ethambutol, Isoniazid, Rifabutin , Rifapentine	Tuberculosis (TB) is a serious illness that mainly affects the lungs. The germs that cause tuberculosis are a type of bacteria. Primary infection is usually
Bronchopulmonary Dysplasia (BPD)	Difficulty breathing, low oxygen levels, blue skin, blue lips, rapid breathing	tracheostomy, surgery, increase calories, nutrition, medications	A form of chronic lung disease that affects premature babies and infants from damage to the lungs caused by mechanical ventilation (respirator) and long-term use of oxygen. Most infants recover, but some may have long-term breathing problems.
Cystic Fibrosis	coughing, mucus, wheezing, infection, inflamed nasal passages, stuffy nose, sinusitis	remove mucus, loosening mucus, nutrition, antibiotics, pancreatic enzymes, bronchodilators, stool softeners, acid reduction,	Cystic fibrosis affects the cells that produce mucus, sweat and digestive juices. These secreted fluids are normally thin and slippery. But in people with CF, a defective gene causes the secretions to become sticky and thick. Instead of acting as lubricants, the secretions plug up tubes, ducts and passageways, especially in the lungs and pancreas.
Pulmonary Fibrosis	shortness of breath, dry cough, fatigue, tiredness, weight loss, aching muscles, aching joints	pirfenidone, Esbriet, nintedanib, Ofev, Nintedanib, anti-acid, oxygen therapy, Pulmonary rehabilitation, nutrition, lung transplant,	Pulmonary fibrosis is a lung disease that occurs when lung tissue becomes damaged and scarred. This thickened, stiff tissue makes it harder for the lungs to work properly. Pulmonary fibrosis worsens over time. Some people can stay stable for a long time, but the condition gets worse faster in others. As it gets worse, people become more and more short of breath.
Sleep Apnea	snoring, gasping for air, dry mouth, headache, insomnia, difficulty sleeping, sleepiness, irritability, hypersomnia	lose weight, quit smoking, sleep position, CPAP, Continuous positive airway pressure, Tissue removal, Jaw repositioning, implants, tissue shrinkage, erve stimulation, oxygen, ventilation, Adaptive servo-ventilation, ASV	Sleep apnea is a potentially serious sleep disorder in which breathing repeatedly stops and starts. You may have sleep apnea if you snore loudly, and you feel tired even after a full night's sleep.
Covid	dry cough, shortness of breath, loss of smell, loss of taste, fatigue, extreme tiredness, upset stomach, vomiting, loose stools, diarrhea, body aches, headache, congestion, runny nose, stuffy nose, fever, chills	Fever reducers, ibuprofen, acetaminophen, cough syrup, pain relievers, Tylenol, Advil, remdesivir, baricitinib, tocilizumab, oxygen, convalescent plasma, ECMO, extracorporeal membrane oxygenation	The virus that causes COVID-19 spreads most commonly through the air in tiny droplets of fluid between people in close contact. Many people with COVID-19 have no symptoms or mild illness. But for older adults and people with certain medical conditions, COVID-19 can lead to the need for care in the hospital or death.
Strep Throat	throat pain, painful swallowing, swollen tonsils, red, spots, swollen lymph nodes, fever headache rash, nausea, vomiting, body aches	antibiotics, symptom relievers, Advil, Motrin, Tylenol, acetaminophen, rest, fluids, drink water, humidifier, avoid irritants	Strep throat is a bacterial infection that can make your throat feel sore and scratchy. Strep throat accounts for only a small portion of sore throats.

Text	Coughing	Wheezing	Dizziness	Fatigue	Chest pain	Stuffy nose	Sore throat	Snoring	Blue skin	Dry cough	Fever	Sweating	Shortness of breath	Difficulty in breathing while sleeping	Low blood pressure	Prouce mucus	Runny nose	Swelling in ankles	Headache	Round Fingertips	Nausea	Unusal lumps	Weight loss	low energy	total
Acute Respiratory Distress Syndrome	0	0	0	0	0	0	0	0	0	0	0	0	0.12	0	0.12	0	0	0	0	0	0	0	0	0	0.24
Asbestosis	0	0	0	0	0.18	0	0	0	0	0	0	0	0.18		0	0	0	0	0	0.18	0		0	0	0.54
Aspergillosis	0	0	0	0	0	0.24	0	0	0	0	0.24	0	0	0	0	0.24	0.24	0	0	0	0	0	0	0	0.96
Asthma	0.18	0.18	0	0	0	0	0	0	0	0	0	0	0.18	0	0	0	0	0	0	0	0	0	0	0	0.54
Bronchiectasis	0.18	0	0	0.18	0.18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.54
bronchiolitis	0	0	0	0	0	0.12	0	0	0	0	0	0	0	0	0	0	0.12	0	0	0	0	0	0	0	0.24
Bronchitis	0.24	0	0	0.24	0	0	0	0	0	0	0.24	0	0	0	0		0	0	0	0	0	0	0.24	0	0.96
Chronic Bronchitis	0.29	0	0	0.29	0.29	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1.45
chronic obstructive pulmonary disease	0	0.41	0	0	0.41	0	0	0	0	0	0	0	0.41	0	0	0	0	0.41	0	0	0	0	0.41	0.41	2.87
Influenza	0	0	0	0	0	0.29	0.29	0	0	0	0	0	0	0	0	0	0.29	0	0.29	0	0	0	0	0.29	1.45
Mesothelioma	0.18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.18	0.18	0	0.54
Pneumonia	0	0.29	0	0.29	0	0	0	0	0	0	0.29	0.29	0	0	0	0	0	0	0	0	0.29	0	0	0	1.45
pneumothorax	0	0	0	0	0.12	0	0	0	0	0	0	0	0.12	0	0	0	0	0	0	0	0	0	0	0	0.24
Pulmonary hypertension	0	0	0.18	0	0	0	0	0	0.18	0	0	0	0.18	0	0	0	0	0	0	0	0	0	0	0	0.54
Respiratory syncytial virus	0	0	0	0	0	0	0	0	0	0.18	0	0	0	0	0	0	0.18	0	0.18	0	0	0	0	0	0.54
Sleep apnea	0	0	0	0	0	0	0	0.12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.12	0.24
Tuberculosis	0.18	0	0	0	0	0	0	0	0	0	0.18	0	0	0	0	0	0	0	0	0	0	0	0	0.18	0.54

[illegible]