4-2022

Unsupervised Learning With Word Embeddings Captures Quiescent Knowledge From COVID-19 And Materials Science Literature

Tasnim H. Gharaibeh
Western Michigan University, tasneemgh@gmail.com

UNSUPERVISED LEARNING WITH WORD EMBEDDINGS CAPTURES QUIESCENT KNOWLEDGE FROM COVID-19 AND MATERIALS SCIENCE LITERATURE

Tasnim Gharaibeh, Ph.D.

Western Michigan University, 2022

Millions of scientific papers are produced each year and the scientific literature is continuing to grow at a head-spinning speed. Thus, massive scientific knowledge exists in solid text, but all these publications make it difficult, if not impossible, for researchers to keep in up to date with discoveries, even within a narrow scientific area. This massive amount of information also makes it difficult to find implicit and hidden connections, relationships, and dependencies within the information that may guide the direction of future research or lead to valuable new insights. So, there is a need for algorithms or models that can scan the text of millions of papers to uncover new scientific knowledge and search for hidden connections within this knowledge. For computer algorithms, to utilize this resource, it should be converted in terms of numbers and represent the words in some mathematical form. This is where artificial intelligence and machine learning can help. Advanced algorithms in machine learning and natural language processing can be used to make large databases more useful and easier to handle by both researchers and clinicians. We used Word2Vec for our implementation and trained many unsupervised word-embedding models on different data sets in materials science and in the medical field to extract hidden knowledge, relations, and interactions based on words that appear in similar contexts in the text while often having similar meanings. So far, we have adopted three main models. The first is trained within additive manufacturing (AM), targeting the powder bed fusion (PBF) processes, such as selective laser sintering (SLS), selective laser melting (SLM), and direct metal laser sintering (DMLS), with the goal of extracting new knowledge to improve AM processes and address material properties depending on the process

used. Other properties inherent to the materials, such as the giant magnetocaloric effect, are also addressed in a specific model. The second model is trained within COVID-19 drugs literature to address what insights can be obtained on candidate drugs to treat COVID-19. Finally, the third model is trained within COVID-19 vaccine literature to predict good candidate vaccines. We thus demonstrate how word embeddings can help extract hidden knowledge from the published literature in very distinct areas of research.

UNSUPERVISED LEARNING WITH WORD EMBEDDINGS CAPTURES QUIESCENT KNOWLEDGE FROM COVID-19 AND MATERIALS SCIENCE LITERATURE

by

Tasnim Gharaibeh

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Computer Science
Western Michigan University
April 2022

Doctoral Committee:

Dr. Elise de Doncker, Chair
Dr. Alvis Fong
Dr. Pnina Ari-Gur

TABLE OF CONTENTS

Table of Contents—Continued

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background and Context

Machine learning approaches are usually divided into three categories: supervised, unsupervised, and reinforcement learning. Supervised learning deals with labeled data, data that contains both the inputs and the desired outputs, to train the model. It uses linear regression, neural networks, nearest neighbor, or naive Bayes algorithms to predict or classify data. Reinforcement learning allows machines or software agents to automatically decide the required behavior within a specific situation to maximize its performance. Using Q-Learning, temporal difference (TD), or deep adversarial networks, reinforcement learning algorithms can be applied to robotic hands, computer-played board games, and self-driving cars. Unsupervised learning is a machine learning technique that allows the model to work on its own to find structure and discover patterns and information that were previously undetected, without the need to supervise the model. It mainly deals with unlabeled data as input [1]. Clustering and association are two types of unsupervised learning. The association rules establish relations among data objects inside large databases, whereas clustering splits the dataset into groups based on their similarities.

The word embedding technique is one of the most important, versatile examples of unsupervised learning models, leading to a type of language modeling or word representation that allows words with similar meanings to be understood by machine learning algorithms [2]. As a representation of words obtained from an unlabeled large corpus, by embedding both semantic and syntactic meanings, word embedding can be represented as:

$$V \rightarrow \mathbb{R}^D : w \mapsto \overrightarrow{w}$$

mapping a word *w* from a vocabulary V to a real-valued vector in an embedding space of dimension D [3]. To compute the similarities between vectors in the embedding space, the cosine similarity function can be used, which is defined as

$$\text{similarity}(w1, w2) = \frac{\overrightarrow{w1}.\overrightarrow{w2}}{\|w1\|\,\|w2\|}$$

According to this function, the nearest neighbors of a word *w* are given as a list of words $v \epsilon V \backslash \{w\}$, sorted in descending order [3].

Word embedding is an important method commonly used in the tasks of modern natural language processing (NLP), such as semantic analysis [4], retrieval of information [5], dependency parsing [6] [7] [8], query answering [9] [10], and computer translation [9] [11] [12]. In these techniques, individual words are represented as real-valued vectors, often comprising tens or hundreds of dimensions, in a predefined vector space. Each word is mapped to one vector. The vector values are learned using an approach that resembles a neural network, and, hence, the technique is often lumped into the field of deep learning. The word-embedding methods vary and can be grouped into different categories. For example, models can be classified as either paradigmatic or syntagmatic models, based on the word distribution information [13] [14]. The most important aspect of the syntagmatic model is that words co-occur in the text region, whereas for the paradigmatic model, it is a similar context that matters. Another grouping that is dependent on how the word embedding is generated divides models into two classes: matrix factorization and sliding-window sampling. The first method is based on the word co-occurrence matrix, where word embeddings are obtained from a matrix decomposition. In sliding-window sampling, data sampled from sliding windows are used to predict the context word [15]. Word embeddings have many important applications, such as idiomaticity analysis, sentiment analysis, syntax analysis, speech (POS) tagging, named entity

recognition, as well as textual entailment [16]. Various word-embedding models are available, such as Word2Vec, GloVe, and FastText.

The Word2Vec method is based on a shallow neural network with two layers, which takes a large corpus of text as its input and produces a vector space, typically of several hundred dimensions, with remarkable linear relationships called analogies. These allow for math operations, such as vec("king") - vec("man") + vec("woman") ≈ vec("queen") [17]. As an unsupervised learning technique, and a machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses, it comprises two techniques: CBOW (Continuous Bag of Words) and Skip-gram. CBOW predicts the probability of a word given a context, while Skip-gram predicts the context given a word. These machine learning models have great benefits and a significant impact in many areas of science and technology, including medical research and materials science.

1.1.1 COVID-19 Drugs

The increase in knowledge and understanding of diseases and drugs is associated with a growth in information. The resulting big data is ready to deliver the benefits of the application of machine learning (ML) in diagnosis, prognosis, and drug development. AI and machine learning models can change drug discovery radically by extracting hidden patterns, structures, relations, and evidence from biomedical data. Pharmaceutical companies have used AI for drug discovery and development.

At a time when COVID-19 patients flood hospitals worldwide, physicians are searching for effective antiviral therapies to save lives. However, there is still a lack of proven effective medications against COVID-19. Multiple vaccine trials and treatments are underway, yet they need more time and testing. Furthermore, the SARS-CoV-2 virus that causes COVID-19 appears to affect various animals in different ways (with respect to infection and spreading), which limits preclinical animal studies. So, one of the many major challenges of the pandemic

has been to find effective treatments for the virus. The rapid spread of coronavirus has created an immediate need for comprehensive therapies, which is a massive undertaking for scientists and drug developers.

Since word embedding in machine learning could speed the research and open further options by detecting existing drugs that could help fight COVID-19, we trained an unsupervised learning model using the Word2Vec algorithm to capture latent knowledge about COVID-19 drugs from the most recent literature, while focusing on the keywords: drugs, biomedical, medication, antibodies, immunity, immunology, and vaccines. Depending on the co-occurrence of the words, since words that occur together tend to have a similar context, we used approved drugs, such as remdesivir, or most promising drugs, such as avigan, atazanavir, dexamethasone, and REGN-COV2, as seeds to our model to find the most similar drugs. As a result, we suggest some possible candidate drugs for clinical investigation or support other ones in different testing phases for approval, according to their high similarity with the currently approved and most promising treatments.

1.1.2 COVID-19 Vaccine

In order to address the COVID-19 pandemic by limiting transmission, an intense global concerns the development of a safe and effective vaccine, which generally requires several years of pre-clinical and clinical stages of evaluation, as well as strict regulatory approvals. However, because of the unprecedented impact of COVID-19 worldwide, the development and testing of new vaccines have been accelerated. There are currently some authorized, not yet approved, vaccines to fight COVID-19, besides other ones in clinical evaluation or in a pre-clinical stage, and many more being researched. Vaccine discovery methods have been costly, and it may take many years to develop an appropriate vaccine against a specified pathogen. In the COVID-19 case, vaccines are essential in reducing morbidity and mortality, especially since the virus establishes itself in the population. Thus, the development of a new vaccine

should be accelerated to save lives. When most people in a community are vaccinated against a disease, the ability of the pathogen to spread is limited, which also provides indirect protection to people who cannot be vaccinated. Machine learning can accelerate the discovery of effective vaccines or suggest components of a vaccine after understanding the viral protein structure and assist medical researchers scouring tens of thousands of related scientific papers [17][18].

1.1.3 Materials Science

Materials science focuses on the relationship between the structure, operation, characteristics, and applications of materials. Current material science research involves many "trial-and-error methods," depending on the experience that guides a large number of experiments, with a limited number of computer simulations and calculations. This conventional method of studying materials is unreliable and consumes time, manpower, materials, and financial resources [19]. Therefore, finding a new efficient research method is necessary to accelerate innovation. Machine learning could speed up the process by constructing models that learn rules from datasets to effectively predict required properties of materials, optimize the process parameters, obtain the hidden relationships among various variables, guide the chemical synthesis route, and improve current material characterization methods [20].

In materials science, additive manufacturing (AM) is an advanced type of manufacturing technique that fabricates parts. It is the process of creating a 3D object through a computer-aided-design (CAD) method by building it layer by layer. Several sectors are now taking advantage of fabricating complex structures using AM technologies to increase functionality, light-weighting, and part number reduction. AM is being used to fabricate end-use products in aircraft, custom parts (e.g., classic car parts, surgical tools), dental restorations, medical implants, automobiles, handling and robotics, and even fashion products [21].

Specific methods can be implemented in additive manufacturing based on different needs to utilize different deposition techniques. Some of these melt the materials and some change the materials into semisolid forms or powders. According to the different heating sources used to change the material states, such as lasers and resistance heaters, additive manufacturing can be divided into seven different processes [22]: vat photopolymerization [23], material jetting, binder jetting, material extrusion, powder bed fusion, sheet lamination, and direct energy deposition [24].

In powder bed fusion (PBF), thermal energy is used, typically in the form of a laser or electron beam. Because of its low-cost quality, this process has widely grown in the last few years and attracted the interest of researchers. PBF also allows for different materials, including metals, plastics, glass, and alloys. Since the powder used in the process can be recycled to produce more parts, recycling is one of the best qualities of PBF [25].

Selective laser sintering (SLS), selective laser melting (SLM), and direct metal laser sintering (DMLS) [24][26] are common versions of powder bed fusion for fabricating metal parts.

SLS is a process of combining powder material (e.g., nylon, thermoplastic, or polystyrene) to build a solid piece layer by layer using heat and pressure, performed with the help of a laser beam. SLS is ideal for biomedical uses, such as prosthetics, pre-surgical planning, and bone scaffolds for tissue engineering [27][28].

SLM is a process that fabricates 3D products layer by layer using high-energy laser beams on a powder bed, where the powder is melted rather than sintered. SLM can process different materials, such as polymers, engineering plastics, and ceramic and metal oxide materials. Furthermore, among all AM processes, SLM is the fastest and is able to fabricate multiple parts in one round [29].

Direct metal laser sintering (DMLS) is also a layer-by-layer, laser-based AM technique, in which the product is developed with the use of metal powders. Basically, the machine fabricates the object on a movable platform by applying incremental layers of the pattern material, with an equal thickness of approximately 0.1 mm [29]. All processes continue, layer by layer, until the object is fabricated completely.

In PBF techniques, process and product development are based on a few factors, such as alloy chemistries, powder characteristics, powder bed processes and beam, powder interactions, and material properties and performance. The wide range of materials that can potentially be processed using PBF is one of the main advantages of PBF techniques. Theoretically, any material that can be melted and resolidified can be processed with PBF; but, in practice, this is not the case. The main materials and alloys used in PBF processes are aluminum, tool steels, titanium, stainless steel, refractory, and super alloys [25].

However, PBF has its challenges, including high material costs; slow speeds; laborious post-processing requirements; and restrictions on materials compatibility. In addition, a variety of defects limits the process in terms of repeatability, precision, and resulting mechanical properties [30], such as porosity, cracks, and residual stresses problems. These challenges can be improved by more research into how the materials interact when being fused together. Also, if more people understand the material properties and how they interact in specific heat (such as lasers) and for different alloys, then there is a better chance to understand how to avoid these defects and make this process more efficient. The world is still waiting for materials that enable the technology to fulfill its true potential.

As an example for these materials and within our area of interest are the materials with compounds exhibiting a giant magnetocaloric effect (GMCE). By applying/removing a magnetic field, the isothermal entropy or the adiabatic temperature will change. A sharp change in magnetization associated with the structural transformation from the high-temperature

austenite phase to the low-temperature martensite phase of lower crystallographic symmetry results in large GMCE [31]. Simply, it is the ability of some magnetic materials to heat up when they are magnetized and cool down when are demagnetized in a thermodynamic cycle. Ni-Mn-based Heusler alloys are one of the materials that exhibit GMCE materials. Heusler alloys (HAs) are materials of enormous interest due to their multifunctional properties including shape memory, magnetoresistance (MR), thermoelectric effect, half-metallicity, spin filtering, and spin injection [31].

To lend a hand in facing these challenges, we trained several unsupervised word-embedding models using Word2Vec on literature involving powder bed fusion (PBF) as the AM process, trying to determine the main additive manufacturing parameters that affect the alloy properties, after the composition is already known or determined. Furthermore, using these models, we intend to search for new high-performance materials or properties that are similar to the ones currently known and could be used to improve AM.

CHAPTER 2

THE PROBLEM ADDRESSED IN THIS DISSERTATION

2.1 Problem Statement

The vast majority of scientific knowledge is published in articles that provide useful knowledge of the interactions and relationships between data objects as viewed by the authors. However, it is difficult to interpret this knowledge either by ordinary machine learning approaches or advanced statistical analysis. For its leverage, that knowledge from the published literature is represented using vectors of words or word embeddings. Without human labeling or any explicit insertion of pre-knowledge, these embeddings can capture hidden or unknown relations existing within very different fields, such as materials science or COVID-19 drug and vaccine research.

2.2 Research Questions

How can word embeddings, using Word2Vec, help to extract hidden knowledge from the published literature?

What candidate drugs can be obtained from word embeddings using Word2Vec algorithms to treat COVID-19 (the study and results reported in [32])?

What candidate vaccines can be obtained from word embeddings using Word2Vec algorithms to reduce COVID-19 cases (the study and results reported in [33])?

What new knowledge on materials and properties, which are similar to those of known materials and properties, can be extracted to improve AM processes?

2.3 Relevance and Importance of the Research

Our COVID-19 model may suggest possible candidate drugs or vaccines for clinical investigation and support other ones in different testing phases for approval, according to their high similarity with currently approved and most promising treatments. We also hope that our

AM models, particularly on SLS, SLM, and DMLS, could add some improvements to AM processes.

CHAPTER 3

LITERATURE REVIEW

3.1 COVID-19 Drugs

3.1.1 Drug Repurposing

Drug repurposing or repositioning is a strategy that uses previously approved drugs to treat newly emerging and challenging diseases, including COVID-19, by extracting hidden patterns and evidence from biomedical data [34]. This technique reduces development timelines and costs, since the safety of these drugs has already been tested in clinical trials for other applications. The probability of failure is lower because, in preclinical models, the repurposed drug has already been shown to be reasonably effective, and subsequent efficacy tests are less likely to fail at least from a safety point of view if the early-stage tests have been completed [35]. Some successful drug repurposing examples are Zidovudine in 1987 to treat HIV/AIDS, which was originally used to treat cancer, and the repurposing of aspirin in 2015 to treat colorectal cancer, even though its regular use had been for analgesia [35].

3.1.2 Drug Repurposing Using Machine Learning

Machine learning is creating a paradigm shift in medicine, starting with research and ending in clinical applications. A variety of machine learning methods are showing their utility for drug discovery and development, such as naive Bayesian, support vector machines, and, more recently, deep neural networks.

Many works discuss the role of machine learning models in precision medicine. Beck et al. [40] developed a hybrid CNN and RNN model called Molecule Transformer-Drug Target Interaction (MT-DTI), a pre-trained deep learning-based drug-target interaction model to predict whether any commercially available antiviral drugs could work for SARS-CoV-2. The authors identified several known antiviral drugs, such as atazanavir, remdesivir, efavirenz,

ritonavir, and dolutegravir, for the potential treatment of SARS-CoV-2 infection. Furthermore, they suggested considering the list of antiviral drugs identified by the MT-DTI model and also found that several antiviral agents, such as Kaletra (lopinavir/ritonavir), could be used for the treatment. According to the associated inhibitory potency, their result showed that the best chemical compound is atazanavir, which is an antiretroviral medication used to treat and prevent human immunodeficiency virus (HIV) infection, with an inhibitory potency of 94.94 nM against the SARS-CoV-2 3C-like proteinase; remdesivir (113.13 nM) came next, then efavirenz (199.17 nM), ritonavir (204.05 nM), and dolutegravir (336.91 nM). In Beck et. al's prediction, lopinavir, ritonavir, and darunavir may also bind to the replication complex components of SARS-CoV-2 with an inhibitory potency with Kd < 1000 nM, in addition to targeting viral proteinases [40].

Another machine learning technique used to discover candidate drugs for SARS-CoV-2 is graph representation learning. This deep-learning technique is used to predict new links and relations between existing approved drugs. The graph contains relationships between different kinds of medical entities (e.g., diseases, drugs, and proteins). Gysi and his team [36] used this technique to develop a model that presented a SARS-CoV-2 case study with 81 potential repurposing candidates identified. According to their prediction, the virus can be found in different tissues, such as the reproductive system, and brain regions. They developed three network-based drug repurposing strategies, relying on their prediction and depending on network proximity, diffusion, and AI-based metrics. Based on their likely efficacy for COVID-19 patients, in addition to aggregating predictions, their strategies rank all approved drugs. Ritonavir, isoniazid, troleandomycin, cilostazol, chloroquine, rifabutin, and flutamide are the first seven highest ranked COVID-19 candidate drugs.

Another team worked with graph representation learning to construct a comprehensive COVID-19 knowledge graph (named CoV-KGE), using a large scientific corpus of 24 million

PubMed publications to create a knowledge graph with 15 million edges across 39 types of relationships connecting pathways, drugs, genes, diseases, proteins, and expressions of genes and proteins. To build this deep-learning methodology, Zeng et al. [37] used Amazon Web Services' computing resources and graph representation learning techniques. As a result, the team identified 41 repurposed drug candidates (including dexamethasone and melatonin) for COVID-19 treatment distributed over three categories, including anti-inflammatory agents, selective estrogen receptor modulators (SERMs), and antiparasitics. For validation, they used the ongoing COVID-19 trial data set as a validation set. CoV-KGE showed high performance in identifying COVID-19 repurposable drugs, as indicated by its large AUROC metric value (AUROC = 0.85). AUROC, the area under the receiver operating characteristic, is a performance metric that can be used to evaluate classification models [37].

Kuusisto and his team [38] introduced an unsupervised learning method using a word-embedding mining approach to find out candidate treatments for SARS-CoV-2. They used the BioWordVec5 model [39], the most recent prebuilt biomedical word embedding available at the time, which performed well on several benchmark tasks. Given that SARS-CoV-2 is a strain of SARS-CoV they used SARS as an approximation reference to SARS-CoV-2 or COVID-19 because BioWordVec was published before SARS-CoV-2 was discovered and has no reference to SARS-CoV-2 or COVID-19 in its vocabulary. To find a vector representation for treatments, they used the analogy approach [17]. Taking advantage of word-embedding structure and its ability to carry semantic meaning, they used three separate drug-to-disease pairs as a seed treatment for the analogies: metformin to diabetes, benazepril to hypertension, and albuterol to asthma. The drug names were extracted from the FDA's approved drug database [41]. To validate their work, they used these seed drug-disease pairs as analogies to find drugs for Alzheimer's, allergies, and cancer. All drugs resulting from each analogy for each validation disease were the drugs with primary indications for this disease (Alzheimer's, allergies, and

cancer) or a similar disease from the same family. Then they used SARS as the query disease and marked every drug that has either been suggested or is currently under investigation for the treatment of SARS-CoV-2. Their results showed 22 of 50 drugs either suggested or under investigation for treatment against SARS-CoV-2 from the metformin-diabetes analogy, 12 of 50 hits from benazepril-hypertension analogy, and 8 from the albuterol-asthma analogy. Amantadine, rimantadine, zanamivir, oseltamivir are some hits common to all analogies.

3.2 COVID-19 Vaccines

3.2.1 Vaccine Development

A pathogen is a bacterium, virus, parasite, or fungus that can cause disease within the body. Each pathogen has its unique subparts. One important factor is the antigen, which causes the formation of antibodies. The antibodies produced in response to the pathogen's antigen are an important part of the immune system. Often, vaccines contain weakened or inactive parts of a particular antigen that train the immune system to recognize and fight pathogens, which greatly reduces the risk of infection. The immune system can safely recognize the pathogen when it is encountered naturally; and by delivering an immunogen, a specific type of antigen elicits an immune response [42].

To fight virus-infected cells, the host immune system produces antibodies on the surface of B-cells or attacks the virus directly using T-cells. To assist T- and B-cells in attacking and binding with invaders, the human leukocyte antigen (HLA) gene encodes MCH-I and MCH-II proteins, which present epitopes as antigenic determinants [43].

To ensure the safety and efficiency of the new vaccine, many actions and steps are taken, such as pre-clinical studies, Phase I clinical trials, Phase II clinical trials, Phase III clinical trials, Phase IV post-marketing surveillance, and human challenge studies. In the pre-clinical studies, the vaccine is tested first in animal studies for efficacy and safety. Then in a Phase I clinical trial, small groups of healthy adult volunteers receive the vaccine to test for

safety. After that, in a Phase II clinical trial, people who have characteristics for whom the new vaccine is intended are given the vaccine. The next step of testing the vaccine is by giving it to thousands of people, which is a Phase III clinical trial. After the vaccine is approved and licensed, many studies are conducted to learn the future effects of the vaccine in the population in the long term [44].

Since the outbreak of the novel coronavirus, different machine learning approaches have been used to predict potential effective vaccines.

3.2.2 Machine Learning and the Development of COVID-19 Vaccines

To accelerate the vaccine design process by predicting the design of a multi-epitope vaccine for COVID-19, Zikun et al. [45] implemented a supervised deep-learning model (DeepVacPred) in silico. From the available SARS-CoV-2 spike protein sequence, their model was able to predict 26 candidate vaccine subunits, by combining the in silico immunoinformatic and deep neural network strategies. In order to construct a multi-epitope vaccine for the SARS-CoV-2 virus, they identified the top 11 from the 26. According to the best 11 candidates, they suggested a 694aa (consisting of 694 amino acid residues) multi-epitope vaccine that contains 16 B-cell epitopes, 82 CTL epitopes, and 89 HTL epitopes, as a promising vaccine to fight the SARS-CoV-2 viral infection. They also concluded that the designed vaccine can be used successfully with the recent RNA mutations of the virus by investigating the SARS-CoV-2 mutations for RNA.

Since the T-cell and B-cell epitopes could be used to produce good vaccines, as well as recognize neutralizing antibodies, Fast et al. [46] tried to identify 2019-nCoV T-cell and B-cell epitopes based on viral protein antigen presentation and antibody binding properties, using computational tools from structural biology and two supervised neural networks (MARIA, a multimodal recurrent neural network [47] and NetMHCpan4, a trained neural network in [48]). Their results show good antigen presentation scores for 405 potential viral peptides, for both

human MHC-I and MHC-II alleles. Their models also identified two potential neutralizing B-cell epitopes close to the 2019-nCoV spike protein receptor-binding domain (440-460 and 494-506). They also concluded that the spike protein for the SARS-CoV-2 virus could be a potential vaccine candidate. This conclusion was the result of their analysis of mutation profiles of 68 viral genomes spread through four continents that showed no mutations are present near the spike protein receptor-binding domain, and 96 coding-change mutations occurred only in regions with good MHC-I presentation scores [46].

By using the same method, the recurrent neural networks (RNN), and a training set consisting of alpha, beta, gamma, and delta coronavirus spike sequences, Crossman simulated the spike protein sequences of coronaviruses for the previously known sequences and tested their characteristics. This simulation could help in the development of vaccine design to predict alternative possible spike sequences that could arise in the future [49].

In another study that introduces a computational framework from several computational approaches, such as reverse vaccinology and immunoinformatics tools with deep learning, Abbasi et al. [50] predicted a potentially suitable protein vaccine candidate against SARS-CoV-2. Reverse vaccinology uses the expressed genomic sequences to find new potential vaccines, whereas immunoinformatics tools, such as PSORTb, FungalRV, SignalP, TargetP, IEDB, BLASTp, ProtParam, Vaxijen, etc., were modified to classify proteins into positive and negative datasets. As a result, one protein (Spike S–Surface Glycoprotein with accession no. QHD43416.1) was shortlisted as a potential vaccine candidate out of the 10 proteins of SARS-CoV-2. Out of 41 B-cell epitope sequences that were predicted using BcePred, only 15 peptides were identified to be highly antigenic as specified by the Vaxijen server, where a peptide "DLCFTNVY" is predicted as the highest-ranking peptide. For T-cell epitopes, they consider peptides with a 100% conservancy rate, so 45 epitopes were identified by the ProPred-I tool to be conserved out of the 46 predicted MHC class-I binding epitopes, where KIADYNYKL was

found with the highest antigenicity score and 90 out of 94 MHC- II allele binding T-cell epitopes predicted by the ProPred tool with VKNKCVNFN as the highest antigenicity score.

3.3 Machine Learning in Materials Science

Due to the enormous numbers of controllable and uncontrollable parameters, the large size of the data acquired, and the difficulties encountered in analyzing it, processing the data in the AM process is a difficult task. Also, manually detecting the defects in AM processes is time-consuming and laborious work. Therefore, there is a need for applying machine learning (ML) techniques for processing the data and detecting defects.

3.3.1 Optimizing AM Process Parameters

Traditionally, to additively manufacture new materials, simulation methods and the design of experiments are used to implement process parameter development and optimization. The physically-based simulation can reveal the underlying mechanism for the formation of specific features during processing. However, it may suffer from discrepancies with experimental results due to simplified assumptions. On the other hand, the design of experiment approaches involves trial-and-error, which is time-consuming and costly for AM. Therefore, to address these challenges in the process optimization of metal AM, many researchers have explored the feasibility of introducing ML approaches [51].

To predict some properties, such as the high cycle fatigue life of laser powder bed fusion of stainless steel 316, Zhang et al. examined the use of a neuro-fuzzy-based machine learning method. For simulating a complex nonlinear input-output environment, they prepared a dataset that consists of fatigue life data for samples with different processing conditions, such as scan speed, laser power, and layer thickness; post-processing treatments, such as annealing and hot isostatic pressing; and cyclic stresses. They developed two models for the processing/post-processing parameters and tensile properties. The models showed a good prediction accuracy

against the test data, even with the diverse fatigue and fracture properties. The team suggested that these models could be used concurrently, for quality assurance at the manufacturing stage and property assessment stage [52].

In another prediction, Shen et al. developed an artificial neural network (ANN) approach for density prediction of SLS parts. Their ANN is a two-layer supervised neural network supplied with SLS process parameters, such as laser power, scan speed, scan spacing, and layer thickness. To collect experimental training and test datasets, an orthogonal experimental method was employed. They concluded that, by applying the ANN approach, the density prediction of SLS parts is rather accurate. Their method does not need to know the precise model, which is an outstanding advantage [53].

Wang et al. used a neural network combined with GA (Genetic Algorithm) techniques to study SLS optimum process parameters, such as layer thickness (0.2 mm), interval time (1 s), laser power (18 W), scanning speed (1,800 mm/s), hatch spacing (0.12mm), work surroundings temperature (93°C), and scanning mode (subarea and direction change scanning), based on the minimum shrinkage ratio. These parameters were obtained by using the genetic algorithm based on the neural network model, so that the part manufactured under the optimal process parameters results in minimum shrinkage. The GA searched the global optimum with reasonable speed for population size = 30. After applying the measurements five times under the same condition layer, they found that the repeatability of the shrinkage ratio was good, between 0.128% and 0.130%, depending on the part number [54].

3.3.2 Detecting Defects

ML has been successfully used in different defect detection scenarios during the AM process. To detect the track defect and predict the printability of materials in SLM for the application in a factory, Chen et al. introduced a neural network (NN) as a supervised machine learning (ML) method. Their method helps in finding the better parameter combinations for

intelligent defect-free printing. A 3D microscope inserted into the selective laser melting method for in-situ measurement was used to track the process. As findings, they classified five types of the printed single tracks according to the measured surface morphologies, and the results were used as target outputs for the ML model. Then, for a quantitative evaluation of the quality of the tracks, they used four evaluation indicators, which are correlated to the surface morphology and important geometrical characteristics of the printed single tracks. Their approach using a backpropagation neural network model successfully predicted the process window for parameters, such as laser power and scan speed, for a TiB2 reinforced AlSi10Mg composite. Furthermore, they confirmed the feasibility of the prediction model by experiment [55].

In another work, Yadav et al. prepared a balanced labeled dataset (with an equal number of "drift" and "no-drift" data points) for training a supervised Support Vector Machine (SVM) classifier, as a classification algorithm that divides the dataset into multiple classes, to detect the drift in the parts [56]. By drift they mean the non-uniformity that results in "hotspots" in the melt pool signatures. These hotspots are the areas of the highest probability to generate real defects in the part. In addition, they are an indication of drift; these are the areas where the intensity of the signal is higher compared to the rest of the layer. To acquire the in-situ data, a co-axial melt pool monitoring (MPM) system installed on a commercial SLM 280HL machine was used without any additional hardware modifications. Specific geometries incorporating the overheating and lack of fusion drift were printed by varying the volumetric energy density and were used to test the trained algorithm. The trained SVM classifier was tested for the overheating of samples and for lack of fusion, which results in internal porosity in the final part and is a challenging detection task. As a result, their model was able to allocate 98.53% of data points with the correct label—"drift" or "no-drift" [56].

Gobert et al. [57] used in-situ layer-wise imaging with a digital single-lens reflex camera to take images of each layer during the build process, under different lighting conditions. They used these images as data for a supervised ML model, using binary classification techniques, i.e., a linear support vector machine (SVM), for defect detection during the L-PBF process. Computed Tomography (CT) scans were used to evaluate the results. In CT scans, discontinuities, e.g., incomplete fusion, porosity, cracks, and inclusions, were identified using automated analysis tools or manual inspection. After training the model, the resulting accuracy of defect detection during the process was up to 85% [57].

Using a convolutional neural deep network, Baumgartl et al. could detect and identify defects during printing processes with an average balanced accuracy of 96.80%. Their model achieves fast compilation and training without the need for powerful hardware, even though it is very small and light in computational costs. They used thermographic images for training and testing. The images were taken in the printing process of H13 steel specimens, during in-situ off-axis monitoring. For the network training of a geometrical shape, an uncritical defect splatter and critical defect delamination were chosen. Also, the model could output a heatmap to help in deciding the type and location of the error. Their model could be used for other defect types, as well as for in-situ defect detection of L-PBF processes, because the model is based on a single source of information, and there is no need to carry out additional evaluations using expensive and time-consuming methods like X-ray or CT [58].

CHAPTER 4

RESEARCH METHODS

4.1 Data Collection and Processing

4.1.1 COVID-19 Drugs Model

We obtained around 18,600 full-text papers and articles primarily focused on COVID-19 in the years 2019 and 2020 from Science Direct at Elsevier. Through Science Direct application programming interfaces (APIs) (https://dev.elsevier.com/), interactive APIs, and with some Python (3.6) code [59] we performed the text-mining process. We searched the Science Direct database using specific search terms as a first step in cleaning the data. These terms combined "COVID-19 OR 2019-nCoV OR COVID19 OR nCoV-19 OR Sars-CoV-2" with keywords, such as drugs, biomedical, medication, antibodies, immunity, immunology, and vaccines.

4.1.2 COVID-19 Vaccines Model

From Science Direct at Elsevier, we obtained around 54,427 full-text published papers and articles primarily focusing on COVID-19 in the years 2019, 2020, and 2021 (see Table 1). In order to perform the text-mining process, we used Science Direct application programming interfaces (APIs) (https://dev.elsevier.com/), interactive APIs, and some Python (3.6) code [59]. We searched the Science Direct database using specific search terms as a first step in cleaning the data. These terms combined "COVID-19 OR 2019-nCoV OR COVID19 OR nCoV-19 OR Sars-CoV-2" with vaccines, biomedical, medication, antibodies, immunity, immunology, and drugs.

Table 1.
The Obtained COVID-19 Data from Science Direct in the Elsevier

	2019	2020	2021
Drugs	**99**	**4870**	**3475**
Vaccines	**110**	**7850**	**5115**
Biomedical	**32**	**1888**	**1469**
Medication	**21**	**4274**	**2685**
Antibodies	**104**	**4497**	**2931**
Immunity	**115**	**7462**	**4764**
Immunology	**35**	**1540**	**1091**
Sum	**516**	**32381**	**21530**
Toltal		**54427**	

4.1.3 AM Model

We obtained around 95,990 full-text papers and articles primarily focused on Ni-Mn-based alloys, SLM, SLS, DML in the years from 2012 to 2022 from Science Direct at Elsevier (see Table 2).

Table 2.
The Obtained SLS, SLM, and DMLS Data from Science Direct at Elsevier

Year	Ni-Mn-based alloys	SLM	SLS	DMLS
2022	64	32	23	19
2021	6800	4076	2313	2453
2020	6443	4054	2309	2331
2019	5574	3368	1947	2095
2018	4935	2909	1629	1820
2017	4422	2222	1234	1446
2016	3697	1805	1045	1278
2015	3697	1459	834	1082
2014	3180	1316	701	972
2013	3045	1173	624	882
2012	2506	968	478	730
Total	44363	23382	13137	15108

Through Science Direct application programming interfaces (APIs) (https://dev.elsevier.com/), interactive APIs, and some Python (3.6) code [59], we performed the text mining process.

4.2 Cleaning Data

Extensive cleaning processes were needed to reduce the noise in the corpus, speed up model training, and improve model efficiency. We deleted words with a length of more than 20 letters; usually, these words are URLs or descriptions related to the metadata. Then we converted all uppercase characters in the corpus to lowercase, so COVID and covid will have one vector. Finally, we deleted the stop words and punctuation from the corpus, so “vaccine” and “vaccine,” will have one vector. We started with a data file of size 820 MB and obtained a reduced file of size 600 MB for the COVID-19 drugs model and from a file of size 2.74 GB to a reduced file of size 2 GB for COVID-19 vaccine. Finally, for the AM model, an initial data file of size 6.58 GB yielded a reduced file of size 4.42GB.

4.3 Word2Vec Training

We used a combination of the Word2Vec implementation in genism [60] and the open-source code implemented by Vahe Tshitoyan et al. [61], with a few modifications to train our models.

Several parameters that affect both training speed and efficiency are recognized by Word2Vec. In our models, we only consider vocabulary that occurs more than five times because words that appear in a billion-word corpus only once or twice or less than five times are probably uninteresting typos and garbage. Furthermore, there is not enough knowledge to include any practical training on those terms, so it is better to ignore them. The vector size, or size of the embedding, in the models is 200. Larger size values need more data from training

but can lead to better (more precise) models. Usually, fair vector size values are in the tens to hundreds[61].

To speed up the training, we set the workers parameter to 16 workers, which relates to the number of threads to employ while training. The models were trained using a 10th generation Intel i7 processor that features 8 Cores and 16 Threads. Since Python contains excellent built-in tools for both multiprocessing and threading, we adjusted the code for Word2Vec to use multiple threads. In the gensim Word2Vec, the input stream of sentences is chunked (fixed number of sentences,100 by default) into jobs that are sent to threads for training [62]. Models use a skip-gram algorithm for the training with negative sampling loss. With negative sampling, we randomly select just a small number of "negative" words (let's say n = 15) for which to update the weights. In this context, a "negative" word is one for which the neural network should output a 0. Negative sampling decreases the computational burden of the training process and increases the efficiency of the resulting word vectors by having each training sample only modify a small percentage of the weights, rather than all of them [63]. It was found that skip-gram with negative sampling loss (n = 15) performs best according to Tshitoyan [61].

With respect to Google code [64], the author indicated that CBOW is faster, while Skip-gram is slower, but the latter does a better job with infrequent words because Skip-gram seeks the prediction of the context given a word instead of the prediction of a word given its context like CBOW does. The window size was 8, which relates to the maximum distance within a sentence between the actual and the predicted word. The initial learning rate was 0.01 and would drop linearly to 0.0001 as training progressed in 30 epochs. The learning rate aims to decide what proportion of the observed error should be used to change weights. The typical range is between 1 and 10^{-6}. The model will oscillate if the learning rate is set too high, and the model will converge too slowly to a solution by setting it too low. A learning rate of 0.01

usually works with most networks [65]. To optimize the computation, we use a hierarchical SoftMax training algorithm, which employs a binary tree to represent all words in the vocabulary [66]. We also trained a model with a different parameters combination according to some other references. For example, we set the number of vector dimensions to 400 instead of 200 [67], the window size of words to 9 instead of 8 [68], the negative samples to 5 instead of 15 [68], and we used CBOW instead of skip-gram [17]. We performed some verification and the results, for example in the similarities, were too far to be considered as similar words to a specific queried word. In addition, several strange words and words without meaning occurred as similar words. So we decided not to consider the model with these parameters

Table 3.
The Top 10 Similarities for COVID-19, Drugs, Fever, and Hygiene

Covid19	Drugs	Fever	Hygiene
covid_19	medications	dry_cough	Handwashing
disease	antiviral_drugs	fever_cough	hand_hygiene
sarscov2	repurposed_ drugs	shortness_breath	personal_hygiene
pandemic	Medicines	dyspnoea	good_hand_hygiene
coronavirus	antiinflammatory_ drugs	symptoms	hygiene_practices
2020	Treatments	malaise	Sanitation
patients	Repurposed	cough_shortness_ breath	preventive_measures
covid	antimalarial_ drugs	fever_fatigue	wear_mask
cases	drug_candidates	flulike_symptoms	Distancing
severe	existing_drugs	nonproductive_ cough	cough_etiquette

CHAPTER 5

RESULTS AND DISCUSSION

5.1 COVID-19 Drugs Model

5.1.1 Verification

Some generally known words for COVID-19 and their top 10 similarities were used to verify our model, CDVec. We chose covid19, drugs, fever, and hygiene since these and their similarities are well-known. Similar words tend to occur together and will have a similar context, and words with similar contexts end up with similar vectors. As preiously shown in Table 3, the top 10 similar words for covid19 are covid_19, disease, sarscov2, pandemic, coronavirus, 2020, patients, covid, cases, and severe. These words are often found in the same context as covid19. The word drugs has medications, antiviral_drugs, repurposed_drugs, medicines, anti-inflammatory_drugs, treatments, repurposed, antimalarial_drugs, drug_candidates, and existing_drugs as similar words. On the other hand, the top 10 similar words for fever are well-known and main symptoms for COVID-19. These words are dry_cough, fever_cough, shortness_breath, dyspnoea, symptoms, malaise, cough_shortness_breath, fever_fatigue, flulike_symptoms, and nonproductive_cough. The top 10 similar words for the fourth word, hygiene, are handwashing, hand_hygiene, personal_hygiene, good_hand_hygiene, hygiene_practices, sanitation, preventive_measures, wear_mask, distancing, and cough_etiquette. These similarities are some of the most important ways to protect against COVID- 19 and many other diseases.

5.1.2 Similarities for Some Drugs

The therapies currently under investigation include drugs that have been used to treat autoimmune diseases, antiviral drugs, and antibodies from people who have recovered from COVID-19.

In this section, we discuss results for some current candidates for COVID-19 drugs that are found in the published literature with their similarities.

Looking for drug similarities reveals more candidate drugs to be tested and other ones to be excluded, or it may increase the chance for under-testing candidates. Without any explicit insertion of prior chemical or structural knowledge about the drugs or COVID-19 treatments, CDVec can find candidate drugs based on word co-occurrences. We divided the results into the categories of antiviral drugs, anti-inflammatory drugs, and antibodies cocktail.

5.1.2.1 Antiviral Drugs

5.1.2.1.1 Remdesivir

The antiviral drug remdesivir (Veklury) was approved in October 2020 by the U.S. Food and Drug Administration (FDA) to treat certain patients hospitalized with COVID-19. Remdesivir is recommended for the treatment of COVID-19 in adults and pediatric patients (12 years of age and older and weighing at least 40 kg) who need hospitalization [69]. However, this approval does not include the entire patient population, and scientists are working hard to develop other effective treatments. CDVec suggests lopinavir-ritonavir, favipiravir, remdesivir_chloroquine, darunavir, lopinavir_ritonavir, hydroxychloroquine, and galidesivir as top candidates with highest similarities as shown in Figure 1.

5.1.2.1.2 Avigan/Favipiravir

Since 2014, Avigan has been approved as an influenza antiviral drug for manufacturing and sale in Japan, as it selectively inhibits the RNA polymerase required for the viral replication of influenza [70], and this is the basis for pursuing approval for treating COVID-19. It has been approved, at least on an emergency basis, for COVID-19 by many nations, such as Russia and India, and is on the way to being formally approved by other countries, such as Japan.

CDVec shows that inhibitors_remdesivir_favipiravir, roseltamivir_umifenovir, boceprevir, rimantadine_amantadine, avipiravir_un_dergone, zalcitabine, (favipiravir_ undergone), ribavirin_ remdesivir, and adefovir are the top similar drugs to Avigan, and that may increase their chance to be used as COVID-19 treatments (Figure 2).

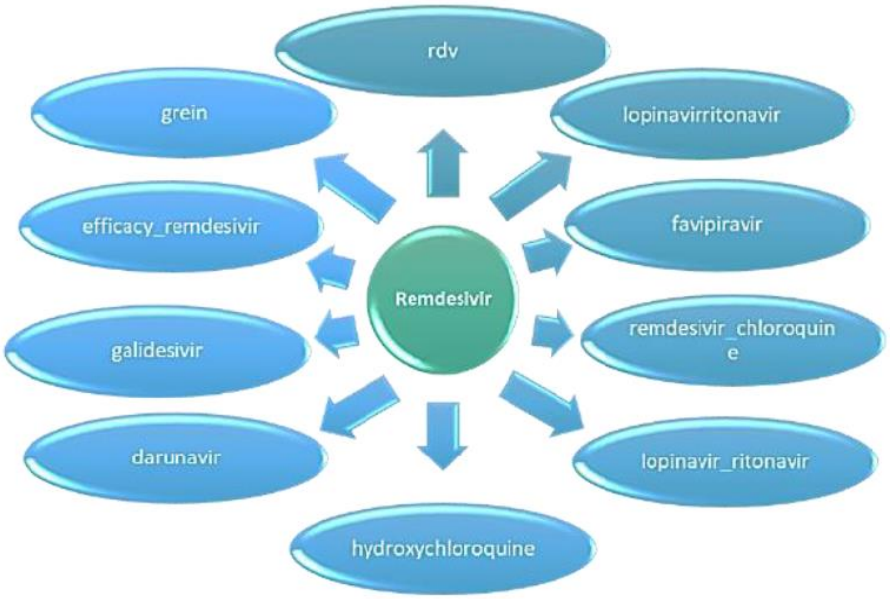

Figure 1. *The Top 10 Similar Words for Remdesivir*

Figure 2. *The Top 10 Similar Words for Avigan*

5.1.2.1.3 Atazanavir

Atazanavir is a prescription drug that was originally approved by the FDA for the treatment of HIV in adults and children from three months of age [71]. Atazanavir belongs to the protease inhibitors class of medications and functions to decrease the amount of HIV in the blood [72]. In a recent study [73], it was found that atazanavir inhibits SARS-CoV-2 replication, in Vero cells and a human pulmonary epithelial cell line, alone or in combination with ritonavir (RTV). Atazanavir/ritonavir also weaken virus-induced enhancement of interleukin 6 (IL-6) and tumor necrosis factor-alpha (TNF-α) levels, and the results clearly suggest that atazanavir should be nominated among the repurposed drugs undergoing clinical trials in the battle against COVID-19. Other studies suggest atazanavir according to the associated inhibitory potency as the best chemical compound followed by remdesivir, efavirenz, ritonavir, and dolutegravir for treating COVID-19. When CDVec obtains the similar words for atazanavir, it finds the same order for remdesivir (with similarity = 0.594158) and efavirenz (with similarity = 0.594158), then ritonavir (with similarity = 0.594398), but with higher similarity for dolutegravir (with similarity = 0.654778).

On the other hand, CDVec indicates more drugs that are similar to atazanavir but with similarities higher than the previous drugs, such that nelfinavir, ritonavir, cobicistat, tipranavir, and saquinavir.

5.1.2.2 Anti-Inflammatory

5.1.2.2.1 Dexamethasone

Dexamethasone, a corticosteroid, is identical to a natural hormone formed by the adrenal gland. When the body does not produce the natural hormone, it replaces it to relieve inflammation, such as redness, swelling, pain, and heat; treat some cancer types; heal intestinal disorders (e.g., colitis, severe allergies, and asthma); treat certain forms of arthritis and skin, blood, kidney, eye, and thyroid diseases [74]. Dexamethasone has decreased mortality in

patients diagnosed with COVID-19 by 28 days among those undergoing intrusive mechanical ventilation or randomized oxygen on their own but not among those without any respiratory assistance [75][76].

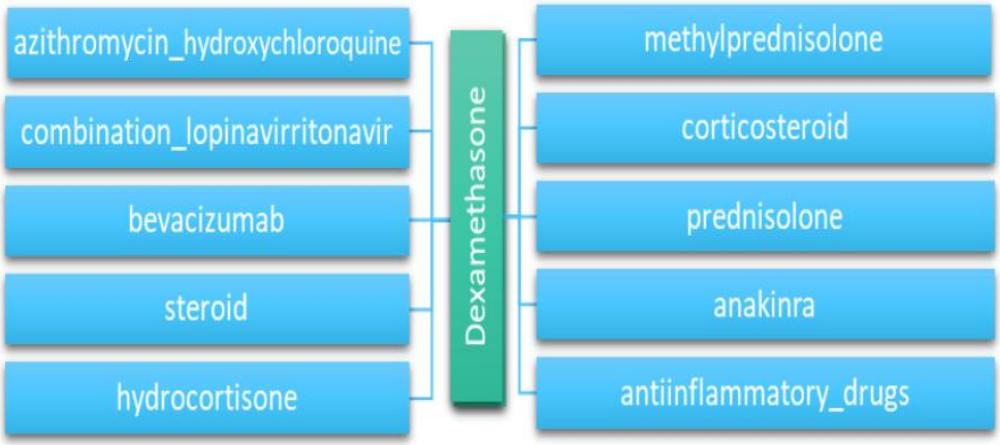

Figure 3. *The Top 10 Similar Words for Dexamethasone*

CDVec proposes the top similar words for dexamethasone as methylprednisolone, corticosteroid, prednisolone, anakinra, bevacizumab, steroid, anti-inflammatory_drugs, combination_lopinavirritonavir (see Figure 3). From these similarities, methylprednisolone, anakinra, bevacizumab, and azithromycin_hydroxychloroquine are promising treatments, so their presence in dexamethasone similarities increases their chances to be considered as future treatments for COVID-19.

5.1.2.3 Antibody Cocktail

An antibody is a protein produced by immune cells in just the right shape and size to bind to a certain foreign object, such as a virus or bacteria in the blood, at a particular point. By entering the antibody, these invaders may be used to directly resist infection or they can be labeled for immune cell death.

5.1.2.3.1 REGN-COV2

The antibody cocktail in REGN-COV2 trials is a combination of two effective, complementary, virus-neutralizing antibodies, monoclonal antibodies (REGN10933 and REGN10987), that originates from a human survivor of COVID-19. Regeneron scientists selected them after evaluating and investigating thousands of options. Each antibody binds to the spike proteins on the virus to block the interaction between the virus and healthy cell, which makes the binding less successful. According to the preclinical studies, REGN-COV2 reduced the virus concentration and associated harm in the lungs [77]. We get many other antibody candidates from our model with high similarities (80%). We obtain ty027, js016, brii196, ctp59, brii198, scta01, 5_vector_expresses, humanized_igg1, nct04348877, and nct04344015 as the top 10 similar words to REGN-COV2 in order (see Figure 4). The first six are antibodies under investigation in different phases as a treatment for COVID-19. Their high similarities to REGN-COV2 make them strong candidates. Furthermore, nct04348877 and nct04344015 are ClinicalTrials.gov identifiers for plasma-rich antibodies from patients recovered from COVID-19.

5.1.3 Similarity Visualization

In Figure 5, we used T-SNE, a machine learning algorithm for data visualization, to reduce the dimensionality of our word-embedding model by mapping the 200-dimensional data in CDVec to two dimensions and visualizing the top 20 similarities between words (remdesivir, fever, drugs, hygiene, avigan, covid19, atazanavir, regncov2, and dexamethasone).

Note the overlap between the similar words for remdesivir, avigan, and dexamethasone, whereas the similar words for atazanavir are separate even though atazanavir is an antiviral medicine like remdesivir and avigan.

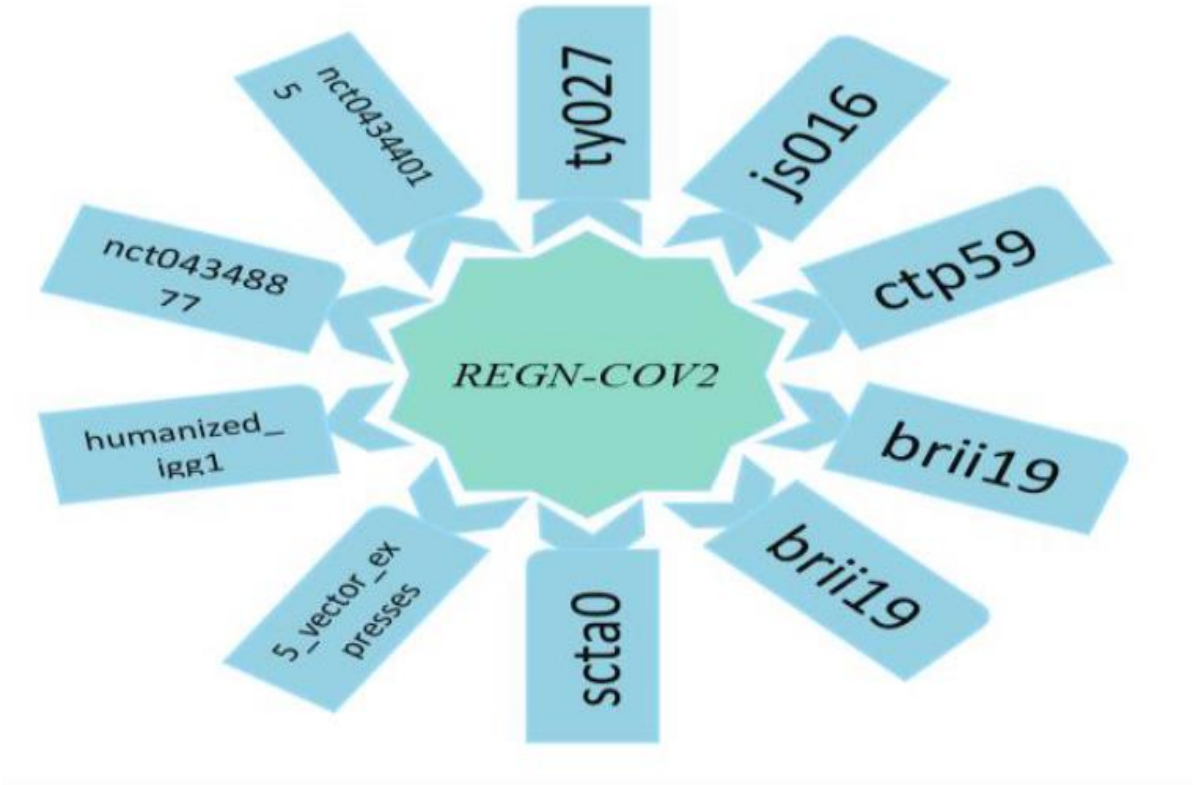

Figure 4. *The Top 10 Similar Words for REGN-COV2*

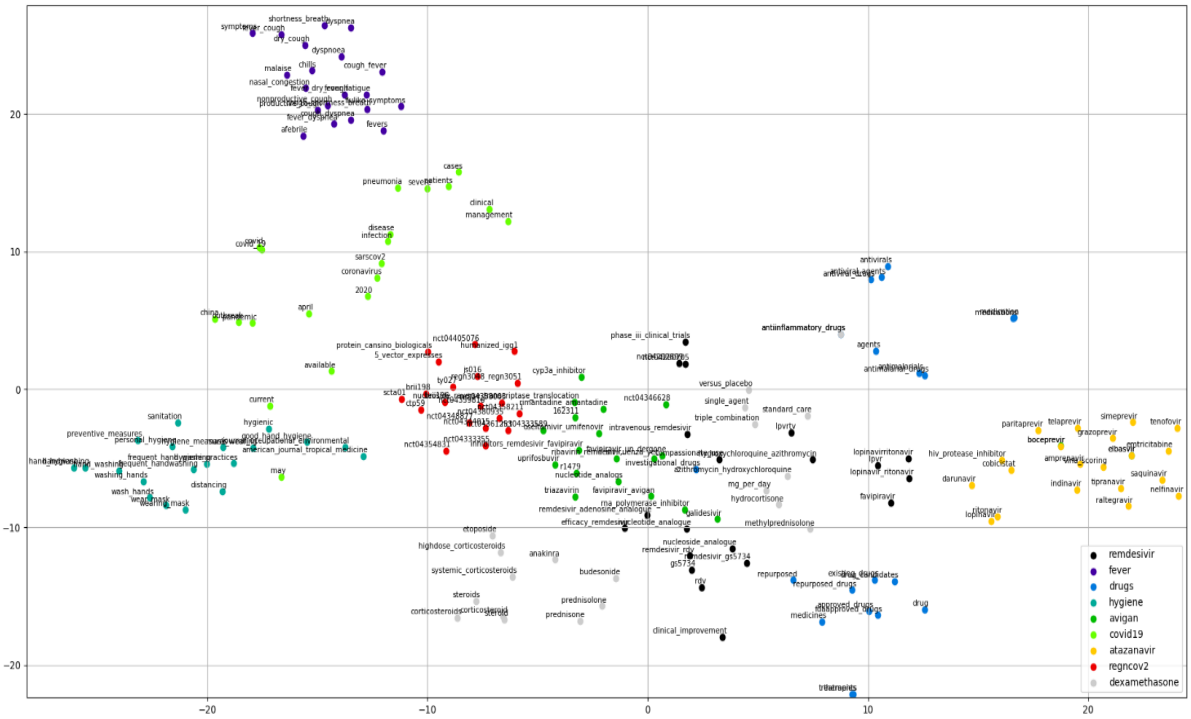

Figure 5. *The Similar Words for Remdesivir, Fever, Drugs, Hygiene, Avigan, COVID-19, Atazanavir, REGN-COV2, and Dexamethasone*

5.2 COVID-19 Vaccine Model

5.2.1 Verification

To verify our COVID-19 vaccine model (CVW2V) results, some general known words and phrases related to COVID-19 vocabulary, such as SARS, vaccines, wearing a mask, and dry cough, were chosen as seeds for the model to test the relationships between these words and their similarities, depending on human judgment as shown in Table 4. For the first word, SARS, the top similar words are sarscov, sarscov1, sars_coronavirus, middle_east_respiratory_syndrome, sarscov2, severe_acute_respiratory_syndrome, mers_sars, corona_virus, infec-tion, and pandemic. The first eight similarities are similar names for COVID-19, whereas infection and pandemic are more general references related to COVID-19. Similar words for vaccines are vaccine_candidates, inactivated_vaccines, mrna_vaccines, immunization, live_attenuated, viral_vector, proteinbased_vaccines, subu-nit_vaccines, rabies_vaccine, and hpv_vaccine. Most of these are types of vaccines. Wearing a mask, as a protective means for COVID-19, has washing_hands, avoid_touching_face, wearing_gloves, frequent_handwashing, maintaining_social_distance, staying_home, face_covering, personal_hygiene, avoiding_crowds, and cough_etiquette as the top 10 similarities, which are also all protective measures against COVID-19. As a symptom of COVID-19, we chose dry cough, which resulted in the top 10 similarities of shortness_breath, fever_cough, dyspnoea, chills, runny_nose, chest_tightness, sore_throat, difficulty_breathing, myalgias, and diarrhea_abdominal_pain. Thus, the similarities found are also symptoms of COVID-19.

5.2.2 Similarities for Some Vaccines

Vaccines are being developed with different technologies, some well-known and others completely new for human vaccines, such as peptide and nucleic acid technologies.

Table 4.
The Top 10 First Similarities for SARS, Vaccines, Wearing a Mask, and Dry Cough

SARS	vaccines	wearing mask	dry cough
sarscov	vaccine_candidates	washing_hands	shortness_breath
sarscov1	inactivated_vaccines	avoid_touching_face	fever_cough
sars_coronavirus	mrna_vaccines	wearing_gloves	Dyspnoea
middle_east_respiratory_syndrome	immunization	frequent_handwashing	Chills
sarscov2	live_attenuated	maintaining_social_distance	runny_nose
severe_acute_respiratory_syndrome	viral_vector	staying_home	chest_tightness
mers_sars	proteinbased_vaccines	face_covering	sore_throat
corona_virus	subunit_vaccines	personal_hygiene	difficulty_breathing
infection	rabies_vaccine	avoiding_crowds	Myalgias
pandemic	hpv_vaccine	cough_etiquette	diarrhea_abdominal_pain

5.2.2.1 Whole Virus Vaccine

These are live-attenuated and inactivated vaccines. Live-attenuated vaccines use a weakened (or attenuated) pathogen that causes disease and creates a strong and long-lasting immune response. Because these vaccines are so close to the real natural infection, only one or a maximum of two doses are needed to give lifetime immunity against the pathogen. These are used to protect against rotavirus, yellow fever, and chickenpox [78]. The inactivated vaccines use the dead version of the pathogen. Therefore, these vaccines do not provide an immunity protection as strong as activated vaccines, so several doses are needed over time to get lasting immunity. They are used to immunize against hepatitis A, flu, polio, and rabies [78]. Sinopharm and CoronaVac are examples of the inactivated virus vaccines for COVID-19.

5.2.2.1.1 Sinopharm

Sinopharm, also known as BBIBP-CorV, was developed by the National Pharmaceutical Group Corporation (CNPGC) in China, with 86% efficacy against the COVID-19 infection [79]. As the most similar words for Sinopharm, CVW2V found sinovac, coronavac, inactivated_vaccine_candidate, ad5ncov, sputnik_v, covaxin, co-dagenix, adenovirusvectored, adults_aged_18–59_years, and wibp_vaccine (see Figure 6).

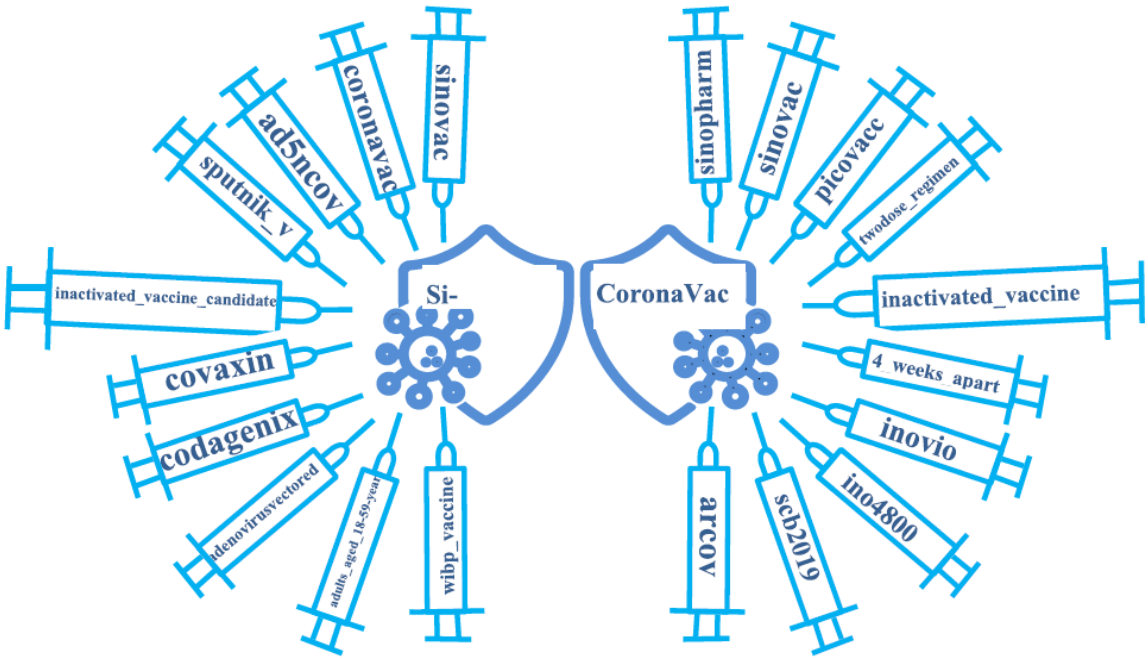

Figure 6. *The Similar Words for the Inactivated Virus Vaccines for COVID-19, Sinopharm, and CoronaVac*

Sinovac is a biopharmaceutical company in China that focuses on the research, development, manufacturing, and commercialization of vaccines. This company developed the CoronaVac inactivated virus COVID-19 vaccine. Ad5-nCoV, or Convidicea, was the first novel coronavirus vaccine for COVID-19 in China and is a viral vector-based vaccine. CoronaVac and Convidicea are under emergency use authorization status in China, Mexico, Pakistan, Hungary, and Chile [80]. Sputnik-V (or Gam-COVID-Vac) is a viral vector vaccine for COVID-19 developed by the Gamaleya Research Institute of Epidemiology and

Microbiology in Russia. Registered on 11 August 2020 by the Russian Ministry of Health and followed by an interim analysis of the trial was published in February 2021, the trial indicated 91.6% efficacy. Over a billion doses of that vaccine were requested by Russia, Argentina, Belarus, Hungary, Serbia, and the United Arab Emirates for immediate distribution globally [81]. Covaxin is another inactivated virus-based COVID-19 vaccine developed in India, by Bharat Biotech in collaboration with the Indian Council of Medical Research. After the second dose, it showed 81% intermediate effectiveness in preventing COVID-19 in those who had not been previously infected [82]. Codagenix is a clinical-stage synthetic biology company that uses programs to recode the genomes of viruses in order to create live-attenuated vaccines or viruses that can be used to prevent viral infections. It generated COVI-VAC, a single-dose, intranasal, live-attenuated vaccine against COVID-19. In Phase I, COVI-VAC showed good results with respect to safety and protection after just one dose in the gold standard animal model and is currently being evaluated in the clinical phase. COVI-VAC is designed to produce immunity against all SARS-CoV-2 proteins to protect against a range of SARS-CoV-2 strains [83]. Adenovirus vectored is a vector-based vaccine that adds a gene for the coronavirus vaccine into a modified version of a chimpanzee adenovirus that can then enter human cells but not replicate inside, so the vaccine can target the spike proteins that SARS-CoV-2 uses to enter human cells [84]. Both the wibp_vaccine and BBIBP-CorV are in-activated virus COVID-19 vaccines developed by Sinopharm [79].

5.2.2.1.2 CoronaVac

CoronaVac is an inactivated virus COVID-19 vaccine developed by the Chinese company Sinovac Biotech. It has been in Phase III clinical trials in several countries. Because it does not need to be frozen, it can be transported and refrigerated at 2–8 °C (36–46 °F), which is the same temperature at which traditional flu vaccines are kept [85]. When CoronaVac was used as a seed in our model, the top similar words were sinopharm, inactivated_vaccine,

sinovac, picovacc, twodose_regimen, 4_weeks_apart, inovio, ino4800, scb2019, and arcov (see Figure 6).

Picovacc is the previous name for CoronaVac, which requires a twodose_regimen and 4_weeks_apart, corresponds to the treatment plan that specifies the amount and schedule for vaccination. Inovio, or ino4800, is a DNA vaccine, INO-4700, against MERS CoV that is currently in preparation to initiate a Phase II vaccine trial [86]. The scb2019 subunit vaccine candidate is developed by Clover Pharmaceuticals [87]. Arcov is the first mRNA vaccine in China and may be ready for final stage trials overseas in May 2021[88].

5.2.2.2 RNA or mRNA Vaccines

mRNA vaccine development uses a new technology, based on proteins, to trigger an immune response. They have several benefits compared to other types of vaccines, including shorter manufacturing times and, because they do not contain a live virus, no risk of causing disease in the person getting vaccinated. Moderna and Pfizer-BioNTech are examples of mRNA vaccines for COVID-19. Under an Emergency Use Authorization (EUA) in December 2019, the FDA authorized emergency use of these vaccines to prevent or reduce the COVID-19 infection in individuals of age 16 years and older. We used these two vaccines as seeds for our model to search for similar candidate vaccines.

5.2.2.2.1 Moderna

The top similarities for Moderna are pfizer_biontech, mrna1273, mrnabased_vaccine, encapsulated_mrna, johnson_johnson, astrazeneca, curevac, arcturus, gsk, and gx19 (see Figure 7).

The mrna1273 similarity is another name for Moderna. CureVac COVID-19 is an mRNA COVID-19 vaccine candidate developed by CureVac N.V. and the Coalition for Epidemic Preparedness Innovations (CEPI), which is different from the current mRNA

COVID-19 vaccines, Pfizer–BioNTech and Moderna. CureVac uses unmodified RNA, while the other two use nucleoside-modified RNA. CureVac is in Phase III clinical trials, as of April 2021 [89]. Arcturus, or ARCT-02,1, is an mRNA COVID-19 vaccine candidate developed by Arcturus Therapeutics and currently in Phase II [90]. Gsk, or VAT00002, is a protein subunit COVID-19 vaccine candidate in Phase II and developed by Sanofi Pasteur and GSK [91]. The gx19 DNA COVID-19 vaccine candidate is developed by Genexine and in Phase II status [92].

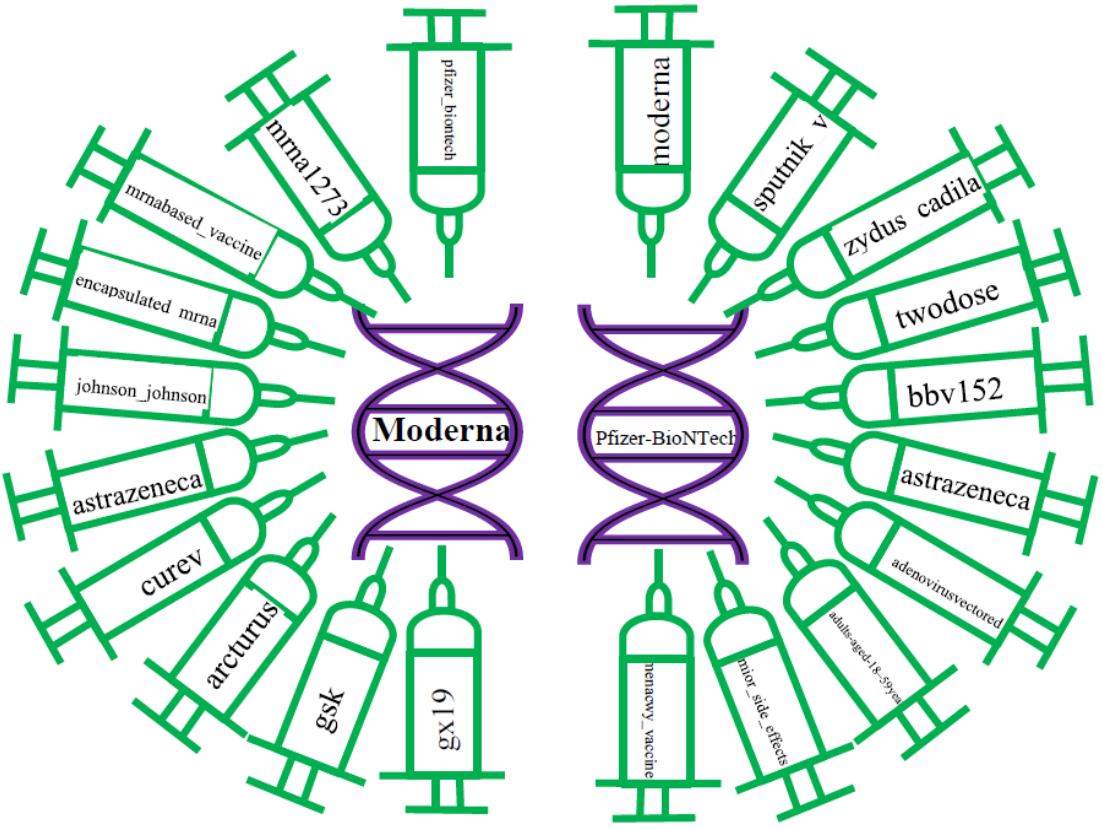

Figure 7. *The Similar Words for Moderna and Pfizer-BioNTech as mRNA Vaccines for COVID-19*

5.2.2.2.2 Pfizer-BioNTech

The similar words found for Pfizer-BioNTech are astrazeneca, moderna, sputnik_v, zydus_cadila, twodose, bbv152, adenovirusvectored, adults_aged_18–59_years, minor_side_effects, and menacwy_vaccine (see Figure 7).

An Indian multinational pharmaceutical company, zydus_cadila, develops a DNA plasmid-based COVID-19 vaccine named ZyCoV-D, which is expected to get emergency use authorization in May or June of 2022 [93]. BBV152, or Covaxin, is an inactivated virus-based COVID-19 vaccine, which was developed by the Indian Council of Medical Research and Bharat Biotech [94]. The MenACWY vaccine is a single injection given into the upper arm that protects against four meningococcal bacteria strains that cause meningitis and blood poisoning: A, C, W, and Y [95].

5.2.2.3 Viral Vector

Viral vector is a type of vaccine that uses a modified version of one virus as a vector. This vector is a DNA molecule used as a vehicle to artificially carry foreign genetic material into another cell. Even though it has two strange objects that enter the body, this type of vaccine does not cause infection with either the virus used as the vector or the source of the antigen. The genetic material it delivers does not integrate into a person's genome [96]. Johnson & Johnson's Janssen and Oxford-AstraZeneca are examples of a viral vector-based COVID-19 vaccine.

5.2.2.3.1 Johnson & Johnson's Janssen

Johnson & Johnson's Janssen is a one dose viral-based COVID-19 vaccine developed by Janssen Vaccines company. We used Johnson & Johnson's Janssen as a seed to our model. Pfizer, astrazeneca, nonreplicating_viral_vector, eli_lilly, merck, bgb_dxp593_beigene, ad26covs1, sclamp, plantderived_vlp, and greffex are some similar words for this vaccine (see Figure 8).

In addition, eli_lilly is an antibody-based treatment for COVID-19 with an Emergency Use Authorization from the Food and Drug Administration [97]. Merck is a pharmaceutical company working on developing an oral medication for COVID-19 and is in a Phase II clinical

trial with promising early findings [98]. BGB-DXP593 is a neutralizing anti-body, or protein-based therapy, under investigation for SARS-CoV-2 in a clinical trial for participants with mild-to-moderate COVID-19 by BeiGene [99]; ad26covs1 is another name for Johnson & Johnson's Janssen vaccine. Sclamp, or V451, is a subunit-based COVID-19 vaccine developed by the University of Queensland and the Australian pharmaceutical company CSL Limited and is still in Phase I [100]. The plantderived_vlp vaccine is a different type of COVID-19 vaccine. Instead of using a viral vector, it uses living plants as bioreactors to produce non-infectious versions of viruses, called Virus-like Particles, or VLPs. This type of vaccine was developed by the Canada-based biopharma Medicago and GlaxoSmithKline and is currently in Phase III clinical trials [101]. Greffex is a vaccine based on an adenovirus viral vector, currently in a pre-clinical phase and developed by Greffex, a genetic engineering company [102].

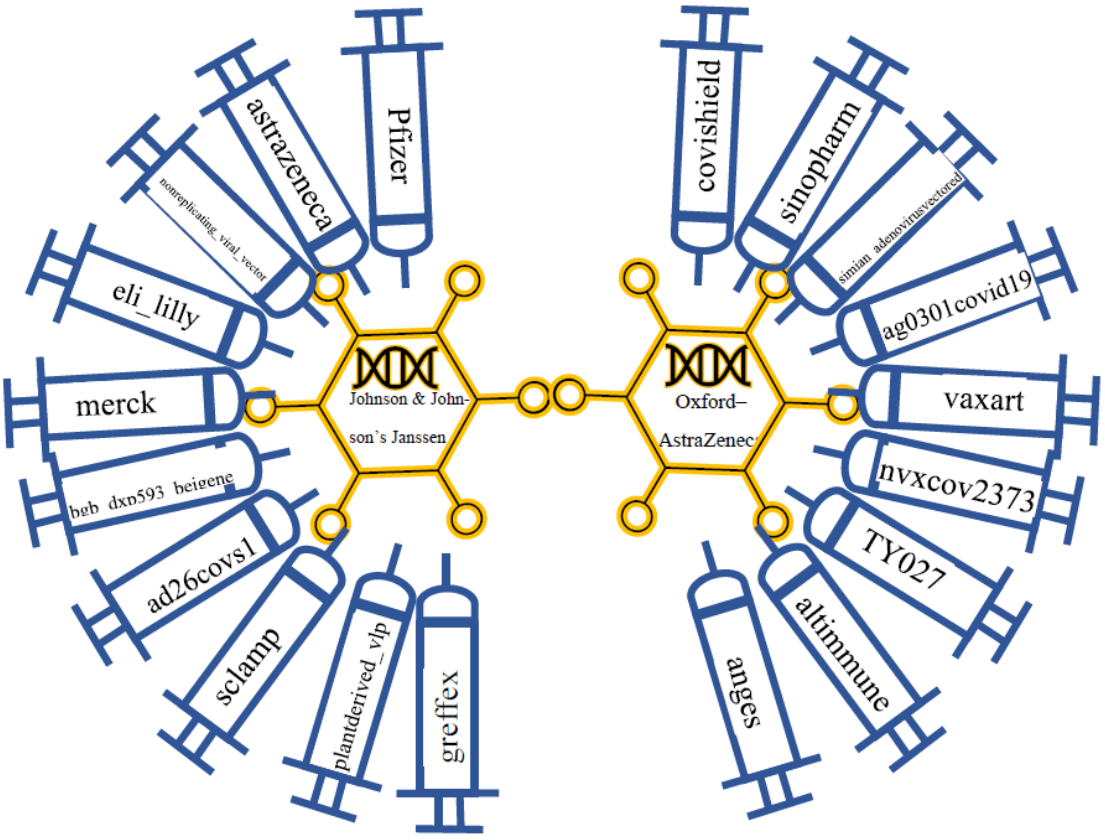

Figure 8. *The Similar Words for Viral Vector-Based COVID-19 Vaccine, Johnson & Johnson's Janssen, and Oxford–AstraZeneca*

5.2.2.3.2 Oxford-AstraZeneca

Oxford-AstraZeneca is a viral vector-based COVID-19 vaccine developed by Oxford University and AstraZeneca. According to CVW2V, interesting similar words for Oxford-AstraZeneca are covishield, sinopharm, anges, ag0301covid19, vaxart, nvxcov2373, simian_adenovirusvectored, TY027, and altimmune (see Figure 8).

The word covishield is another name for Oxford-AstraZeneca. Anges, or AG0302-COVID-19, is a DNA candidate vaccine developed by AnGes Inc. in Japan and currently in Phase II/III [103]. Vaxart is a nonreplicating viral vector oral vaccine platform for COVID-19 in Phase I [104]; and nvxcov2373 is a protein-based COVID-19 vaccine [105]. The simian_adenovirusvectored vaccine is used as a vector in the vector-based vaccine instead of the human adenovirus because of its advantages over the human adenovirus. TY027 is a treatment for patients with COVID-19 to slow the progression of the disease and accelerate recovery and has the potential to provide temporary protection against infection with SARS-CoV-2 [106]. Altimmune is a single-dose viral vector COVID-19 vaccine that triggers a broad immune response—neutralizing IgG, mucosal IgA, and T-cells—and is still in Phase I [107].

5.3 AM Model

5.3.1 Verification

We used some general keywords related to AM processes, parameters or alloys, such as SLM (Selective Laser Melting), LPBF (Laser Powder Bed Fusion), hatch spacing, binder saturation level, NiMn, and powder bed, as seeds to our AM model (AMW2V), to find their most similar words. The results are shown in Table 5.

Table 5.
The Similar Words for SLM, LPBF, Hatch Spacing, Binder Saturation Level, NiMn, and Powder Bed

SLM	LPBF	hatch spacing	binder saturation level	NiMn	powder bed
lpbf	slm	Hatch_distance	zb56	coni	Powderbased
ebm	ebm	scan_spacing	layer_thickness_binder_saturation	Mncu	pbf_processes
dmls	dmls	laser_spot_diameter	air_gap_raster_angle	feni	loose_powder
sebm	lbm	line_spacing	orientation_raster_angle+	fecu	build_platform
lmd	waam	008mm	air_gap_raster_width	nico	Recoating
alsi10mg	ss316l	layer_thickness_kept_constant	wc12_co	comn	Sls
ti64	scan_strategy	laser_power_200w	zcast	Mnfe	unsintered_powder
ss316l	174_ph	131w	farzadi_m_solatihashjin	ni3fe	blown_powder
ti6al4v	in625	focal_position	exone_innovent	Cozn	Layering
alsi12	in718	stripe_width	beam_speed_hatch	feco	binder_jetting

For SLM, the most similar words were lpbf, ebm, dmls, sebm, lmd, alsi10mg, ti64, ss316l, ti6al4v, and alsi12. The lpbf (laser powder bed fusion), ebm (Electron Beam Melting),

dmls (Direct Metal Laser Sintering), sebm (Selective Electron Beam Melting), and lmd (Laser Metal Deposition) are different AM techniques for SLM, whereas alsi10mg, ti64, ss316l, ti6al4v, and alsi12 are popular alloys used in AM. AlSi10Mg is an aluminum alloy commonly used in a powder form in the AM process because of its distinctive and convenient properties, such as good hardness, low density, dynamic toughness, high corrosion resistance, and high mechanical strength of the end parts [108]. Ti64 (or Ti-6Al-4V) is a well-known light titanium alloy used in AM with characteristics that make it ideal for many high-performance applications in medical, aerospace, and automotive areas. The high strength, hardness, high plasticity, long fatigue life for the parts, low density, excellent corrosion resistance, and superior biocompatibility with human tissue are the most important characteristics that make it suitable for such applications [109]. SS316L (or 316L stainless steel) is another ideal alloy, with a nickel and chromium content, manufactured by AM due to its properties, such as ease of fabrication, biocompatibility, reasonable cost, sufficient mechanical strength, and corrosion resistance [110]. The AlSi12 alloy is one of the most widely used metal powders in selective laser melting (SLM), which is because of its relatively low price, low melting point, good thermal properties, needed corrosion resistance, and high tensile strength with low specific gravity [111].

The similar words for the lpbf (Laser Powder Bed Fusion) were slm, ebm, dmls, lbm, waam, ss316l, scan_strategy, 174_ph, in625, and in718. The slm (Selective Laser Melting), ebm (Electron Beam Melting), dmls (Direct Metal Laser Sintering), lbm (Laser Beam Melting), and waam (Wire Arc Additive Manufacturing) are different techniques in AM. The ss316l and 174_ph alloys belong to the stainless-steel family of powders that have been used in AM because of their optimal characteristics, whereas the in625 (or Inconel 625) and in718 (or Inconel 718) alloys are members of the nickel alloy family of powders that AM has used to

fabricate many parts for different applications, such as gas turbines for aerospace and energy industries and ship building.

The words hatch_distance, scan_spacing, laser_spot_diameter, line_spacing, 008mm, layer_thickness_kept_constant, laser_power_200w, 131w, focal_position, and stripe_width were the most similar words for hatch spacing. The hatch spacing (or hatch distance) is one of the AM process parameters, particularly laser and scan parameters. It is the separation of two consecutive laser beams, measured by the distance from the center of one beam to the center of the next beam.

Laser_spot_diameter, line_spacing, layer_thickness_kept_constant, laser_power_ 200w, focal_position, and stripe_width are also parameters that influence the properties of parts fabricated by AM processes [112], whereas 008mm and 131w are some values related to the previous parameters.

For the binder saturation level, the most similar words were zb56, layer_thickness_binder_saturation, air_gap_raster_angle, orientation_raster_angle, air_gap_ raster_width, wc12_co, zcast, farzadi_m_solatihashjin, exone_innovent, and beam_speed_hatch. The binder saturation level is the amount of binder deposited through the printhead in the AM printing process. The layer_thickness, air_gap_raster_angle, orientation_raster_angle, air_gap_raster_width, and beam_speed_hatch are some important parameters in the printing process. The printing parameters are the settings and tuning that need to be provided to the process in order to produce the AM product [113]. The words zcast and zb56 are related to an AM process; zcast is a process developed by the Z Corporation that can generate mould tools for the direct printing of complex moulds quickly and inexpensively with a proprietary mould material (ZCast501), and zb56 is a binder in Zcast powder that gives hardiness [114]. WC12Co is a powder consisting of tungsten carbide-cobalt that produces hard, abrasive, and corrosion-resistant coatings, provides service in environments up to 900°F, and

has been used in SLM as a powder to fabricate parts [115]. The name farzadi_m_solatihashjin is given after the names of two professors, A. Farzadi and Mehran Solati-Hashjin, in the Department of Materials and Metallurgical Engineering at the Amirkabir University of Technology, Tehran. Their research areas of interest are biomaterials, 3D printing and bioprinting, with a number of publications and citations. The ExOne Innovent is a 3D printer made by the German manufacturer ExOne for industrial uses.

On the other hand, the similar words for NiMn (Nickel Manganese alloy) were coni (Cobalt Nickel superalloy), mncu (Manganese-Copper alloy), feni (Iron–Nickel alloy), fecu (Copper Iron alloy), nico (Cobalt-Nickel superalloy), comn (Cobalt Manganese alloy), mnfe (Iron Manganese alloy), ni3fe (Nickel Iron alloy), cozn (Cobalt Zinc alloy), and feco (Iron Cobalt alloy). All of these alloys have been used, sometimes in combinations, in AM fabrication parts, especially in LPBF [116][117][118].

The words that had a vector similar to the powder bed vector were powderbased, bf_processes, loose_powder, build_platform, recoating, sls (Selective Laser Sintering), unsintered_powder, blown_powder, layering, and binder_jetting. Powder bed fusion (powderbased or bf_processes) is a technique to fabricate parts in AM using lasers and the powder of alloys. The terms loose_powder and unsintered_powder refer to the extra or unused powder remaining from the powder bed fusion process, which could be reused in a new fabrication. The word build_platform is the part of the printer where the object is printed. The recoating is where a new layer of powder is spread on top of the printed layer in the PBF process. The blown powder process (also known as Directed Energy Deposition (DED) or laser cladding) involves inserting metal powder into a heat source—for example, a laser that melts metal particles together when deposited [119]; binder_jetting is another type to fabricate parts in AM that does not use heat during the materials fusing process.

5.3.2 Similarities for Some Alloys

There are many produced alloys with wanted and required properties that have been used in SLM, SLS, and MDSL. We are interested in Heusler alloys with Giant Magnetocaloric Effect (GMCE) as observed in some of the Ni-Mn-X-based Heusler alloys. The MCE is a property of some magnetic materials that, when heated up, they are placed in a magnetic field (or magnetized) and cooled down when they are removed (or demagnetized). When the applied magnetic field is increased, these materials transform from a low-symmetry martensitic phase to a highly symmetric ferromagnetic (FM), austenitic phase. In some cases, such as with Ni-Mn-based materials, a Giant Magnetocaloric Effect (GME) is exhibited, which is a magnetocaloric effect much larger than MCE in other magnetic materials and can be exploited for efficient cooling. Magnetic cooling could be a clean energy solution to replace conventional vapor compression refrigeration in the future.

The net cooling effect of GMCE can be dominated by many factors, such as the magneto-structural transition [120] and magnetic entropy changes. The magneto structural phase transition is a coupling between structural and magnetic transitions that are exhibited by magnetic materials [121]. The magnetic entropy change causes a temperature change of magnetic materials in an adiabatic process under an external magnetic field.

The SLM process uses the powder of Ni-Mn-Ga to produce it in the form of melt-spun precursors. The resulting material is capable of martensitic transformation and has a uniform chemical composition. This alloy shows a typical plate-like martensitic microstructure. Apart from martensite plates, the microstructure contains a high density of stacking faults and dislocations. From the magnetic susceptibility and magnetization measurements, it turns out that the ternary Ni-Mn-Ga shows a stronger magnetic response than the other alloys [122]. In order to search for more Ni-Mn-based Heusler alloys candidates, we used main words, such as giant magnetocaloric effect, Ni-Mn-based Heusler alloys, magnetic entropy change, and

magneto structural transition as seeds to our model to find some new candidate alloys. We queried our model, AMW2V, to find the most similar words for these key words. Because this resulted in many different alloys, properties, and parameters, we decided to tighten the result by making the key words more specific. We studied some intersections between the word similarities, looking for alloys that occur simultaneously within the similar words for both "giant magnetocaloric effect" and "Ni-Mn-based Heusler alloys." As shown in Figure 9, mnfepge_compounds, ni-mn-in-co, fe49rh51, nicomnsn, mncoge, ni50mn35in15_heusler, nimnsn, ni43mn46sn11, lafesi13based, mnas1-xsbx, gd5ge2si2, mnassb, tb5si2ge2, erco2, mnnige, mnfe2psi, gd5si2ge2, nazn13type, mnfep1xasx, ni-co-mn-in, mnfepsi, ni-co-mn-sb, and nimnz are in the intersection of similarities considered. This introduces them as predicted candidates to be produced in LBPF processes.

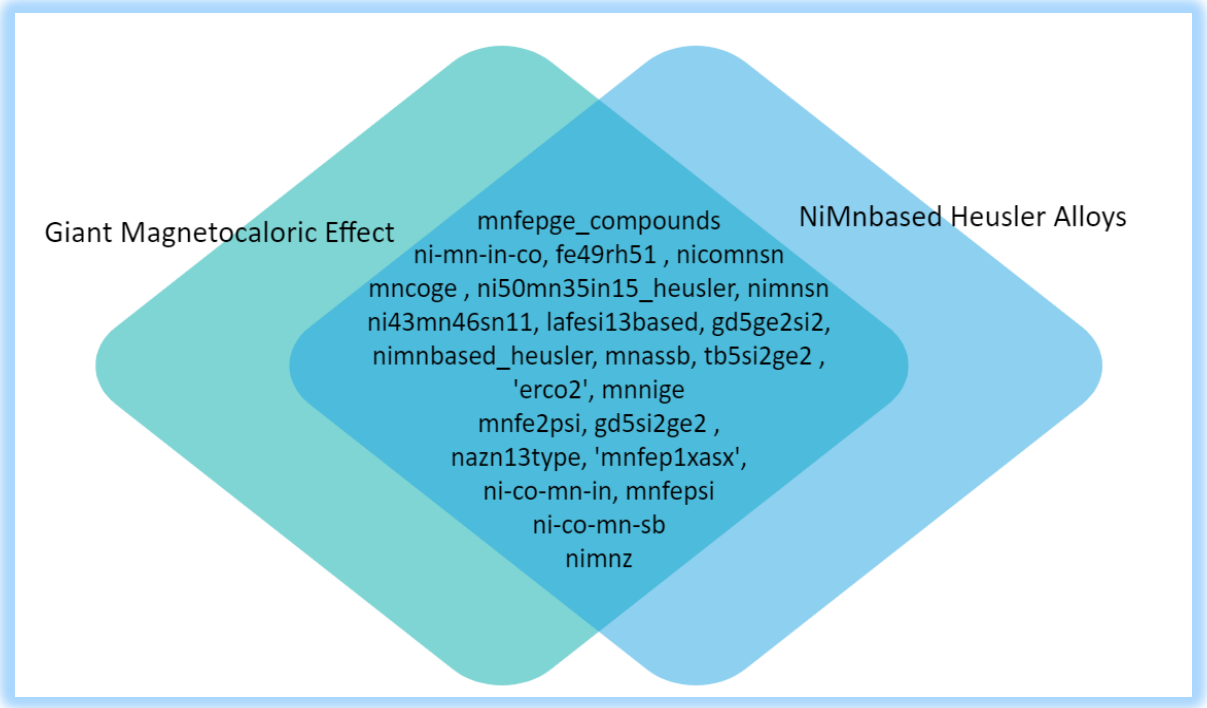

Figure 9. *The Intersection of the Similarities for Both "Giant Magnetocaloric Effect" and "Ni-Mn-Based Heusler Alloys"*

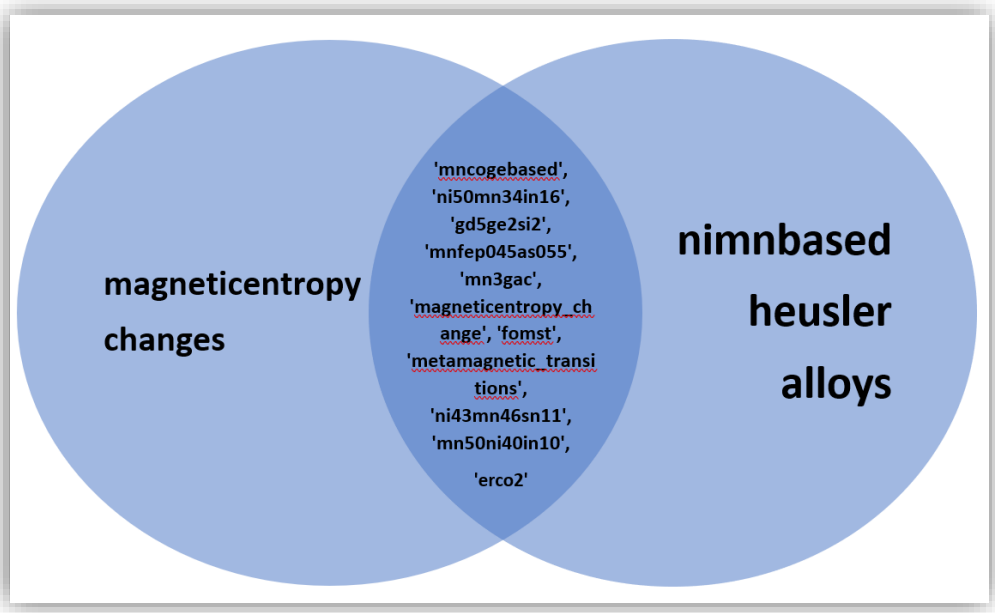

Figure 10. *The Intersection of the Similarities for Both Magnetic Entropy Change and Ni-Mn-Based Heusler Alloys*

Some other intersections between the similarities were obtained, such as the intersection of the similar words between "magnetic entropy change" and "nimnbased heusler alloys" (see Figure 10), and the intersection of the similar words between "magnetic entropy change" and "giant magnetocaloric effect" (see Figure 11).

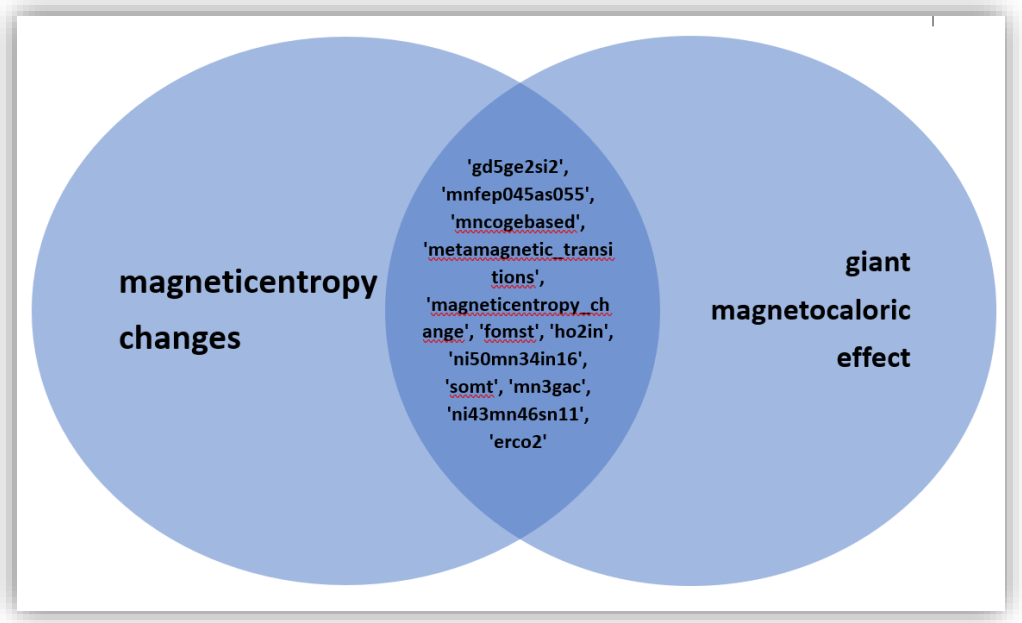

Figure 11. *The Intersection of the Similarities for Both Magnetic Entropy Change and Giant Magnetocaloric Effect*

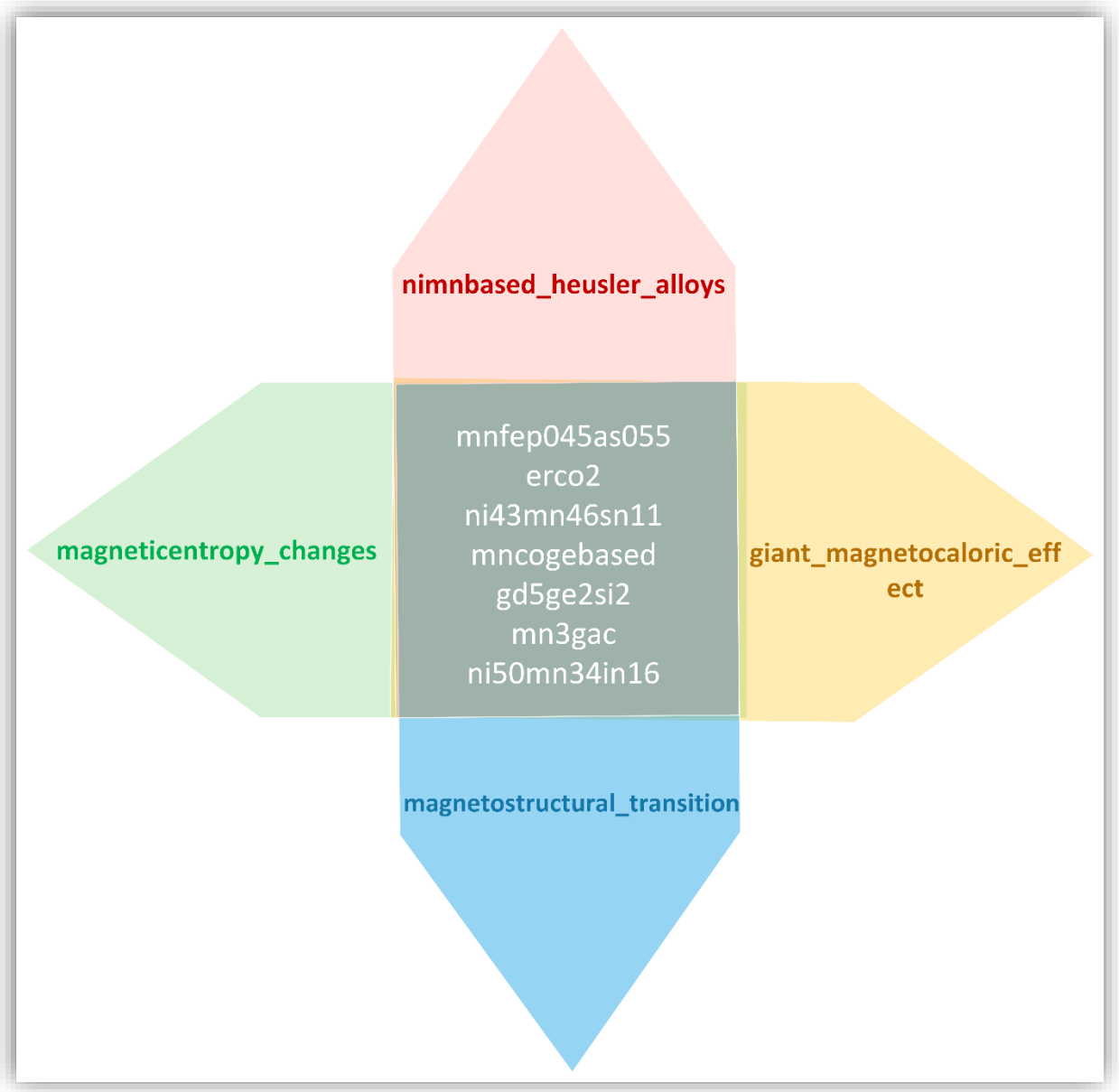

Figure 12. *The Intersection Between Ni-Mn-Based Heusler Alloys, Magnetic Entropy Changes, GMCE, and Magnetostructural Transition*

To find the top candidate alloys, the intersection between the most dominant factors in the GMCE was obtained (see Figure 12). Because mnfep045as055, erco2, ni43mn46sn11, mncogebased, gd5ge2si2, mn3gac, and ni50mn34in16 are the alloys that occur within the similar words for Ni-Mn-based Heusler alloys, magnetic entropy changes, GMCE, and magnetostructural transition, they could be considered as top candidates to be produced in LPBF processes. Alloy MnFeP045As055, as compared to well-known magnetic alloys, such

as gadolinium alloy, is thought to be a promising magnetic material with improved magnetic properties [123]. ErCo2 shows a large magnetocaloric effect, suggesting a high potential for a working substance of magnetic refrigeration at 30-50 K [124]. In addition, ni43mn46sn11 and ni50mn34in16 are alloys with a behavior that allows the magnetic-field-driven transition from martensitic phase with low magnetization to austenite with high magnetization, which results in a GMCE property [125]. MnCoGe-based alloys have a strong interplay between structure and magnetism, which results in the exhibited GME [126]. We also found that gd5ge2si2 is a very good alloy for application as an active regenerator material in room temperature magnetic refrigerators because of the GME with a transition temperature at around 276 K [127]. All of these alloys have not been produced by LPBF processes (as of the date of writing). Thus, they may be good candidate materials with GMCE to be produced by LPBF. However, the usage of As and Gd would be a significant concern in terms of the material's safety [128][129][130].

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

In this work, we trained several unsupervised learning models using word embeddings and the Word2Vec algorithm with information from the COVID-19 literature, especially concerning drugs and vaccines, and with Additive Manufacturing literature focusing on LPBF processes, GMCE, and Ni-Mn-based Huesler alloys.

To answer the research question—*"What candidate drugs can be obtained from word embeddings using Word2Vec algorithms to treat COVID-19?"*—we trained an unsupervised learning model (CDVec) using word embedding and the Word2Vec algorithm to capture latent knowledge about COVID-19 treatments from the most recent literature, while focusing on drugs, biomedicine, medication, antibodies, immunity, immunology, and vaccines. Without any prior explicitly inserted knowledge about drugs, treatment, or biomedical information about COVID-19, our model CDVec was able to indicate some candidate drugs for COVID-19. Depending on the co-occurrence of the words, as similar words that occur together tend to have a similar context, we used approved drugs, like remdesivir, or most promising drugs, like avigan, atazanavir, dexamethasone, and REGN-COV2, as seeds to our model to find the most similar drugs. As a result, CDVec suggests many possible candidate drugs for clinical investigation and supports other ones in different testing phases for approval, according to their high similarity with currently approved and most promising treatments [32].

In addition to our work in on drug discovery using Word2vec for COVID-19 treatments, for the research question—*"What candidate vaccines can be obtained from word embeddings using Word2Vec algorithms to reduce COVID-19 cases?"*—we trained an unsupervised word embedding model (CVW2V) to search for potential candidate vaccines in the COVID-19 literature, according to their similarities with authorized or most promising

existing ones. These studies rely on the property that words sharing similar surrounding words are semantically close. After verifying CVW2V with the terms SARS, vaccines, wearing a mask, and dry cough as general words and their similarities, CVW2V was provided with currently authorized vaccine names, such as Sinopharm and CoronaVac, as inactivated whole virus-based vaccines, Moderna and Pfizer-BioNTech as mRNA vaccines, and Johnson & Johnson's Janssen and Oxford–AstraZeneca as viral vector-based COVID-19 vaccines. As a result, without any explicitly inserted knowledge about COVID-19 or vaccines or any super-vision on the words or objects and their meanings, the CVW2V model was able to suggest several potential target vaccines for clinical investigation and support other ones in different clinical experimental phases, according to their high similarity with currently authorized and most promising vaccines [33].

In order to answer the research question—*"What new knowledge on materials and properties that are similar to those of known materials and properties can be extracted to improve AM processes?"*—we trained an unsupervised word-embedding model (AMW2V), using Word2Vec primarily focused on Ni-Mn-based alloys, SLM, SLS, and DML, in the years from 2012 to 2022 from Science Direct at Elsevier, to search for candidate alloys with GMCE to be produced with LPBF processes. AMW2V predicted some candidates, such as mnfep045as055, erco2, ni43mn46sn11, mncogebased, gd5ge2si2, mn3gac, and ni50mn34in16, according to their similarities with given Ni-Mn-based Huesler alloys exhibiting GMCE properties based on magnetic entropy changes and magnetostructural transition.

6.2. Future Work

For future work in the COVID-19 field, we are planning to obtain more hidden knowledge from the updated published literature about, for example, COVID-19 preexisting conditions, healthcare, vaccine deniers, and possible new strains of COVID-19.

On the materials science side, our research supports a direction of producing candidate alloys using LPBF with optimum process parameter combinations, including laser power, scanning speed, hatch spacing, and layer thickness, allowing manufacturers to accurately produce parts with complex shapes and intricate features from the candidate alloys.

REFERENCES

[1] M. Usama et al., "Unsupervised Machine Learning for Networking: Techniques, Applications, and Research Challenges", IEEE Access, vol. 7, pp. 65579-65615, 2019. Available: 10.1109/access.2019.2916648 [Accessed 4 November 2020].Li, Yang & Yang, Tao. (2017). Word Embedding for Understanding Natural Language: A Survey. 10.1007/978-3-319-53817-4.

[2] L. Yu, J. Wang, K. R. Lai and X. Zhang, "Refining Word Embeddings Using Intensity Scores for Sentiment Analysis," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 26, no. 3, pp. 671-681, March 2018, DOI: 10.1109/TASLP.2017.2788182.H. Sch¨utze, C. D. Manning, and P. Raghavan, Introduction to information retrieval. Cambridge University Press, 2008, vol. 39.

[3] T. Schnabel, I. Labutov, D. Mimno, and T. Joachims, “Evaluation methods for unsupervised word embeddings,” Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015.

[4] L.-C. Yu, J. Wang, K. R. Lai, and X. Zhang, “Refining word embeddings using intensity scores for sentiment analysis,” IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 26, no. 3, pp. 671–681, 2018.

[5] H. Schutze, C. D. Manning, and P. Raghavan, Introduction to information retrieval. Cambridge University Press, 2008, vol. 39.

[6] W. Chen, M. Zhang and Y. Zhang, "Distributed Feature Representations for Dependency Parsing," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 3, pp. 451-460, March 2015, DOI: 10.1109/TASLP.2014.2365359.

[7] H. Ouchi, K. Duh, H. Shindo, and Y. Matsumoto, “Transition-based dependency parsing exploiting super tags,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 11, pp. 2059–2068, 2016.

[8] M. Shen, D. Kawahara, and S. Kurohashi, “Dependency parse reranking with rich subtree features,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 7, pp. 1208–1218, 2014.

[9] G. Zhou, Z. Xie, T. He, J. Zhao, and X. T. Hu, “Learning the multilingual translation representations for question retrieval in community question answering via non-negative matrix factorization,” IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 24, no. 7, pp. 1305–1314, 2016

[10] Y. Hao, Y. Zhang, K. Liu, S. He, Z. Liu, H. Wu, and J. Zhao, “An end-to-end model for question answering over knowledge base with cross-attention combining global knowledge,” in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), vol. 1, 2017, pp. 221–231.

[11] B. Zhang, D. Xiong, J. Su, and H. Duan, “A context-aware recurrent encoder for neural machine translation,” IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 25, no. 12, pp.2424–2432, 2017.

[12] K. Chen, T. Zhao, M. Yang, L. Liu, A. Tamura, R. Wang, M. Utiyama, and E. Sumita, “A neural approach to source dependence based context model for statistical machine translation,” IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 26, no. 2, pp.266–280, 2018

[13] F. Sun, J. Guo, Y. Lan, J. Xu, and X. Cheng, "Learning Word Representations by Jointly Modeling Syntagmatic and Paradigmatic Relations", Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on

Natural Language Processing (Volume 1: Long Papers), 2015. Available: 10.3115/v1/p15-1014.

[14] S. Lai, K. Liu, S. He, and J. Zhao, "How to Generate a Good Word Embedding," in IEEE Intelligent Systems, vol. 31, no. 6, pp. 5-14, Nov.-Dec. 2016, DOI: 10.1109/MIS.2016.45.

[15] W. Yin and H. Schütze, "Discriminative Phrase Embedding for Paraphrase Identification", Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2015. Available: 10.3115/v1/n15-1154.

[16] Y. Li and T. Yang, "Word Embedding for Understanding Natural Language: A Survey", Studies in Big Data, pp. 83-104, 2017. Available: 10.1007/978-3-319-53817-4_4.

[17] Mikolov T, Chen K, Corrado G, et al.: Efficient estimation of word representations in vector space. arXiv preprint arXiv: 1301.3781. 2013.

[18] A. Zhavoronkov, Y. A. Ivanenkov, A. Aliper, M. S. Veselov, V. A. Aladinskiy, A. V. Aladinskaya, V. A. Terentiev, D. A. Polykovskiy, M. D. Kuznetsov, A. Asadulaev, Y. Volkov, A. Zholus, R. R. Shayakhmetov, A. Zhebrak, L. I. Minaeva, B. A. Zagribelnyy, L. H. Lee, R. Soll, D. Madge, L. Xing, T. Guo, and A. Aspuru-Guzik, “Deep learning enables rapid identification of potent DDR1 kinase inhibitors,” Nature Biotechnology, vol. 37, no. 9, pp. 1038–1040, 2019.

[19] Y. Liu, T. Zhao, W. Ju, and S. Shi, “Materials discovery and design using machine learning,” Journal of Materiomics, vol. 3, no. 3, pp. 159–177, 2017.

[20] W. Sha, Y. Guo, Q. Yuan, S. Tang, X. Zhang, S. Lu, X. Guo, Y.-C. Cao, and S. Cheng, “Artificial Intelligence to Power the Future of Materials Science and Engineering,” Advanced Intelligent Systems, vol. 2, no. 4, p. 1900143, 2020.

[21] D. Delgado Camacho, P. Clayton, W. J. O'Brien, C. Seepersad, M. Juenger, R. Ferron, and S. Salamone, “Applications of additive manufacturing in the construction industry – A forward-looking review,” Automation in Construction, vol. 89, pp. 110–119, 2018.

[22] ISO/ASTM, 52900-15, Standard Terminology for Additive Manufacturing - General Principles - Terminology, ASTM International, West Conshohocken, 2015, http://dx.doi.org/10.1520/ISOASTM52900-15.

[23] P. Wu, J. Wang, X. Wang, A critical review of the use of 3-D printing in the construction industry, Autom. Constr. 68 (2016) 21–31, http://dx.doi.org/10.1016/j.autcon.2016.04.005.

[24] I. Gibson, D.W. Rosen, B. Stucker, Additive Manufacturing Technologies, 238 Springer, New York, 2010, http://dx.doi.org/10.1007/978-1-4939-2113-3 ISBN 978-1-4939-2112-6.

[25] R. Singh, A. Gupta, O. Tripathi, S. Srivastava, B. Singh, A. Awasthi, S. K. Rajput, P. Sonia, P. Singhal, and K. K. Saxena, “Powder bed fusion process in additive manufacturing: An overview,” Materials Today: Proceedings, vol. 26, pp. 3058–3070, 2020.

[26] Amir Dehghanghadikolaei (2018) Additive Manufacturing as A New Technique of Fabrication. Journal of 3D Printing and Applications - 1(1):3-4.

[27] M. Galati, L. Iuliano, A literature review of powder-based electron beam melting focusing on numerical simulations, Addit. Manuf. (2018), https://doi.org/10.1016/j.addma.2017.11.001.

[28] P. Mercelis, J.P. Kruth, Residual stresses in selective laser sintering and selective laser melting, Rapid Prototyp. J. (2006), https://doi.org/10.1108/13552540610707013.

[29] M. Agarwala, D. Bourell, J. Beaman, H. Marcus, and J. Barlow, “Direct selective laser sintering of metals,” Rapid Prototyping Journal, vol. 1, no. 1, pp. 26–36, 1995.

[30] E. Malekipour and H. El-Mounayri, “Common defects and contributing parameters in powder bed fusion AM process and their classification for online monitoring and control: a

review," The International Journal of Advanced Manufacturing Technology, vol. 95, no. 1-4, pp. 527–550, 2017.

[31] A. Ahmad, S. Mitra, S. K. Srivastava, and A. K. Das, "Giant magnetocaloric effect in Co2FeAl Heusler alloy nanoparticles," Journal of Physics D: Applied Physics, vol. 54, no. 38, p. 385001, 2021.

[32] T. Gharaibeh, E. de Doncker," Unsupervised Learning with Word Embeddings Captures Quiescent Knowledge from COVID-19 Drugs Literature" 2020. Published in The 2020 International Conference on Computational Science and Computational Intelligence.

[33] T. Gharaibeh and E. de Doncker, "Unsupervised Learning Model to Uncover Hidden Knowledge from COVID-19 Vaccines Literature," Computational Science and Its Applications – ICCSA 2021, pp. 544–559, 2021.

[34] S. Mohanty, M. Harun AI Rashid, M. Mridul, C. Mohanty and S. Swayamsiddha, "Application of Artificial Intelligence in COVID-19 drug repurposing", Diabetes & Metabolic Syndrome: Clinical Research & Reviews, vol. 14, no. 5, pp. 1027-1031, 2020. Available: 10.1016/j.dsx.2020.06.068.

[35] S. Pushpakom et al., "Drug repurposing: progress, challenges and recommendations", Nature Reviews Drug Discovery, vol. 18, no. 1, pp. 41-58, 2018. Available: 10.1038/nrd.2018.168.

[36] Gysi, Deisy Morselli et al. "Network Medicine Framework for Identifying Drug Repurposing Opportunities for COVID-19." ArXiv arXiv:2004.07229v1. 15 Apr. 2020 Preprint.

[37] X. Zeng et al., "Repurpose Open Data to Discover Therapeutics for COVID-19 Using Deep Learning", Journal of Proteome Research, 2020. Available: 10.1021/acs.jproteome.0c00316.

[38] F. Kuusisto, D. Page and R. Stewart, "Word embedding mining for SARS-CoV-2 and COVID-19 drug repurposing", F1000Research, vol. 9, p. 585, 2020. Available: 10.12688/f1000research.24271.1.

[39] Y. Zhang, Q. Chen, Z. Yang, H. Lin and Z. Lu, "BioWordVec, improving biomedical word embeddings with subword information and MeSH", Scientific Data, vol. 6, no. 1, 2019. Available: 10.1038/s41597-019-0055-0.

[40] B. Beck, B. Shin, Y. Choi, S. Park, and K. Kang, "Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model", Computational and Structural Biotechnology Journal, vol. 18, pp. 784-790, 2020. Available: 10.1016/j.csbj.2020.03.025.

[41] Federal Drug Administration: Drugs@FDA Data Files. 2020.

[42] J. Mellet and M. S. Pepper, “A covid-19 vaccine: Big strides come with big challenges,” Vaccines, vol. 9, no. 1, p. 39, 2021.

[43] A. Keshavarzi Arshadi, J. Webb, M. Salem, E. Cruz, S. Calad-Thomson, N. Ghadirian, J. Collins, E. Diez-Cecilia, B. Kelly, H. Goodarzi, and J. S. Yuan, “Artificial Intelligence for covid-19 drug discovery and vaccine development,” Frontiers in Artificial Intelligence, vol. 3, 2020.

[44] Center for Biologics Evaluation and Research, “Vaccine development – 101,” U.S. Food and Drug Administration. [Online]. Available: https://www.fda.gov/vaccines-blood-biologics/development-approval-process-cber/vaccine-development-101. [Accessed: 30-Dec-2021].

[45] Z. Yang, P. Bogdan, and S. Nazarian, “An in silico deep learning approach to multi-epitope vaccine design: A SARS-COV-2 case study,” Scientific Reports, vol. 11, no. 1, 2021.

[46] E. Fast, R. B. Altman, and B. Chen, “Potential T-cell and B-cell epitopes of 2019-ncov,” 2020.

[47] B. Chen, M. S. Khodadoust, N. Olsson, L. E. Wagar, E. Fast, C. L. Liu, Y. Muftuoglu, B. J. Sworder, M. Diehn, R. Levy, M. M. Davis, J. E. Elias, R. B. Altman, and A. A. Alizadeh, “Predicting HLA class II antigen presentation through Integrated Deep Learning,” Nature Biotechnology, vol. 37, no. 11, pp. 1332–1343, 2019.

[48] V. Jurtz, S. Paul, M. Andreatta, P. Marcatili, B. Peters, and M. Nielsen, “NetMHCpan-4.0: Improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data,” The Journal of Immunology, vol. 199, no. 9, pp. 3360–3368, 2017.

[49] L. C. Crossman, “Leveraging deep learning to simulate coronavirus spike proteins has the potential to predict future zoonotic sequences,” 2020.

[50] B. A. Abbasi, D. Saraf, T. Sharma, R. Sinha, S. Singh, P. Gupta, S. Sood, A. Gupta, and kamal rawal, “Identification of vaccine targets & design of vaccine against SARS-COV-2 coronavirus using computational and deep learning-based approaches.,” 2020.

[51] C. Wang, X. P. Tan, S. B. Tor, and C. S. Lim, “Machine learning in additive manufacturing: State-of-the-art and perspectives,” Additive Manufacturing, vol. 36, p. 101538, 2020.

[52] M. Zhang, C.-N. Sun, X. Zhang, P. C. Goh, J. Wei, D. Hardacre, and H. Li, “High cycle fatigue life prediction of laser additive manufactured stainless steel: A machine learning approach,” International Journal of Fatigue, vol. 128, p. 105194, 2019.

[53] F. Yin, J. Wang, and C. Guo (Eds.): ISNN 2004, LNCS 3174, pp. 832–840, 2004.

[54] W. Rong-Ji, L. Xin-hua, W. Qing-ding, and W. Lingling, “Optimizing process parameters for selective laser sintering based on neural network and genetic algorithm,” The International Journal of Advanced Manufacturing Technology, vol. 42, no. 11-12, pp. 1035–1042, 2008.

[55] Y. Chen, H. Wang, Y. Wu, and H. Wang, “Predicting the Printability in Selective Laser Melting with a Supervised Machine Learning Method,” Materials, vol. 13, no. 22, p. 5063, 2020.

[56] P. Yadav, O. Rigo, C. Arvieu, E. Le Guen, and E. Lacoste, “Drift Detection in Selective Laser Melting (SLM) Using a Machine Learning Approach,” Industrializing Additive Manufacturing, pp. 177–191, 2020.

[57] C. Gobert, E. W. Reutzel, J. Petrich, A. R. Nassar, and S. Phoha, “Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging.,” Additive Manufacturing, vol. 21, pp. 517–528, 2018

[58] H. Baumgartl, J. Tomas, R. Buettner, and M. Merkel, “A deep learning-based model for defect detection in laser-powder bed fusion using in-situ thermographic monitoring,” Progress in Additive Manufacturing, vol. 5, no. 3, pp. 277–285, 2020.

[59] “Build software better, together”, [Online]. Available: github.com/ElsevierDev/elsapy.. [Accessed: 1-Nov.-2020].

[60] P. Sojka, “Gensim: Topic Modelling For Humans”, [Online]. Available: radimrehurek.com/gensim_3.8.3/index.html. [Accessed: 4-Nov.-2020].

[61] V. Tshitoyan et al., "Unsupervised word embeddings capture latent knowledge from materials science literature", Nature, vol. 571, no. 7763, pp. 95-98, 2019. Available: 10.1038/s41586-019-1335-8

[62] Sumit, Radim, M. Pinches, C. Moody, Elizabeth, L. Zhang, Guillaume, H. Testing, and Harvy, “Parallelizing word2vec in Python,” Pragmatic Machine Learning, 04-Oct-2013. [Online]. Available: https://rare-technologies.com/parallelizing-word2vec-in-python/. [Accessed: 21-Feb-2022].

[63] Goldberg, Y. and Levy, O., “word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method”, arXiv e-prints, 2014.

[64] "Google Code Archive - Long-term storage for Google Code Project Hosting". code.google.com. Retrieved 22 October 2020.

[65] Strandqvist, W. (2016). Neural Networks for Part-of-Speech Tagging (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-129296.

[66] Rong, Xin. “Word2vec Parameter Learning Explained.” ArXiv Preprint ArXiv:1411.2738, 2014.

[67] Adewumi, Tosin P., Foteini Liwicki, and Marcus Liwicki. "Word2Vec: Optimal Hyper-Parameters and Their Impact on NLP Downstream Tasks." arXiv preprint arXiv:2003.11645 (2020).

[68] E. Altszyler, S. Ribeiro, M. Sigman, and D. Fernández Slezak, “The interpretation of dream meaning: Resolving ambiguity using latent semantic analysis in a small corpus of text,” Consciousness and Cognition, vol. 56, pp. 178–187, 2017.

[69] “FDA Approves First Treatment for COVID-19”, 22-Oct.-2020. [Online]. available: www.fda.gov/news-events/pressannouncements/fda-pproves-first-treatment-covid-19. [Accessed: 1-Nov.-2020].

[70] “FUJIFILM Toyama Chemical Co. Submits Application for Avigan (Favipiravir) Approval for COVID-19 to Japan's Health Authority”, Trial Site News. 19-Oct.-2020. [Online]. Available: www.trialsitenews.com/fujifilm-toyama-chemical-co-submits

application-for-avigan-favipiravir-approval-for-covid-19-to-japanshealth-authority/. [Accessed: 1-Nov.-2020]

[71] “Atazanavir”, ClinicalInfo. [Online]. Available: clinicalinfo.hiv.gov/en/drugs/atazanavir/patient. [Accessed: 2-Nov.-2020]. [72] “Atazanavir: MedlinePlus Drug Information”, 28-Sept.-2020. [Online]. Available: medlineplus.gov/druginfo/meds/a603019.html. [Accessed: 2-Nov.-2020].

[73] N. Fintelman-Rodrigues et al., "Atazanavir, Alone or in Combination with Ritonavir, Inhibits SARS-CoV-2 Replication and Proinflammatory Cytokine Production", Antimicrobial Agents and Chemotherapy, vol.64, no. 10, 2020. Available: 10.1128/aac.00825-20.

[74] “Dexamethasone: MedlinePlus Drug Information”, 28-Sept.-2020. [Online]. Available: medlineplus.gov/druginfo/meds/a682792.html. [Accessed: 2-Nov.-2020].

[75] "Dexamethasone in Hospitalized Patients with Covid-19 — Preliminary Report", New England Journal of Medicine, 2020. Available: 10.1056/nejmoa2021436 [Accessed 2 November 2020].

[76] M. Matthay and B. Thompson, "Dexamethasone in hospitalised patients with COVID-19: addressing uncertainties", The Lancet Respiratory Medicine, 2020. Available: 10.1016/s2213-2600(20)30503-8 [Accessed 2 November 2020].

[77] U. D. Health, “What is an antibody cocktail and what trials are being done at UC Davis Health?”, [Online]. Available: health.ucdavis.edu/health-news/newsroom/what-is-an-antibodycocktail-and-what-trials-are-being-done-at-uc-davis-health/2020/10. [Accessed: 2-Nov.-2020].

[78] “The different types of covid-19 vaccines,” World Health Organization. [Online]. Available: https://www.who.int/news-room/feature-stories/detail/the-race-for-a-covid-19-vaccine-explained. [Accessed: 02-Jan-2022].

[79] S.-lee Wee and A. Qin, “China approves covid-19 vaccine as it moves to inoculate Millions,” The New York Times, 30-Dec-2020. [Online]. Available: https://www.nytimes.com/2020/12/30/business/china-vaccine.html. [Accessed: 02-Jan-2022].

[80] J. Corum and C. Zimmer, “How the Sinovac Vaccine Works,” The New York Times, 24-Dec-2020. [Online]. Available: https://www.nytimes.com/interactive/2020/health/sinovac-covid-19-vaccine.html. [Accessed: 02-Jan-2022].

[81] D. Y. Logunov, I. V. Dolzhikova, D. V. Shcheblyakov, A. I. Tukhvatulin, O. V. Zubkova, A. S. Dzharullaeva, A. V. Kovyrshina, N. L. Lubenets, D. M. Grousova, A. S. Erokhova, A. G. Botikov, F. M. Izhaeva, O. Popova, T. A. Ozharovskaya, I. B. Esmagambetov, I. A. Favorskaya, D. I. Zrelkin, D. V. Voronina, D. N. Shcherbinin, A. S. Semikhin, Y. V. Simakova, E. A. Tokarskaya, D. A. Egorova, M. M. Shmarov, N. A. Nikitenko, V. A. Gushchin, E. A. Smolyarchuk, S. K. Zyryanov, S. V. Borisevich, B. S. Naroditsky, and A. L. Gintsburg, “Safety and efficacy of an RAD26 and RAD5 vector-based heterologous prime-boost COVID-19 vaccine: An interim analysis of a randomised controlled phase 3 trial in Russia,” The Lancet, vol. 397, no. 10275, pp. 671–681, 2021.

[82] “COVAXIN® - India's First Indigenous COVID-19 Vaccine,” Bharat Biotech-Vaccines & Bio-Therapeutics Manufacturer in India. [Online]. Available: https://www.bharatbiotech.com/covaxin.html. [Accessed: 04-Jan-2022].

[83] “Codagenix Home,” Codagenix. [Online]. Available: https://codagenix.com/. [Accessed: 04-Jan-2022].

[84] D. S. S. Dutta, “What are adenovirus-based vaccines?,” News, 10-Mar-2021. [Online]. Available: https://www.news-medical.net/health/What-are-Adenovirus-Based-Vaccines.aspx. [Accessed: 04-Jan-2022].

[85] J. Corum and C. Zimmer, “How the Sinovac Vaccine Works,” The New York Times, 24-Dec-2020. [Online]. Available: https://www.nytimes.com/interactive/2020/health/sinovac-covid-19-vaccine.html. [Accessed: 04-Jan-2022].

[86] K. Modjarrad et al., “Safety and immunogenicity of an anti-Middle East respiratory syndrome coronavirus DNA vaccine: a phase 1, open-label, single-arm, dose-escalation trial”, Lancet Infect. Dis., vol 19, no 9, bll 1013–1022, Sep 2019.

[87] “Clover SCB-2019 COVID-19 vaccine,” Welcome to Precision Vaccinations. [Online]. Available: https://www.precisionvaccinations.com/vaccines/clover-scb-2019-covid-19-vaccine. [Accessed: 04-Jan-2022].

[88] Person, “China's mrna COVID-19 vaccine may start late-stage trial in may - state media,” Reuters, 13-Apr-2021. [Online]. Available: https://www.reuters.com/business/healthcare-pharmaceuticals/chinas-mrna-covid-19-vaccine-may-start-late-stage-trial-may-state-media-2021-04-13/. [Accessed: 04-Jan-2022].

[89] “Celonic and Curevac announce agreement to manufacture over 100 million doses of CureVac's COVID-19 vaccine candidate, CVnCoV,” CureVac, 30-Mar-2021. [Online]. Available: https://www.curevac.com/en/2021/03/30/celonic-and-curevac-announce-agreement-to-manufacture-over-100-million-doses-of-curevacs-covid-19-vaccine-candidate-cvncov/. [Accessed: 05-Jan-2022].

[90] “Ascending dose study of investigational SARS-COV-2 vaccine ARCT-021 in healthy adult subjects - full text view,” Full Text View - ClinicalTrials.gov. [Online]. Available: https://www.clinicaltrials.gov/ct2/show/NCT04480957. [Accessed: 05-Jan-2022].

[91] “Izda 2020: PDF: Complement system: Vaccines,” Scribd. [Online]. Available: https://scr.toolsminati.com/document/505230578/izda2020. [Accessed: 05-Jan-2022].

[92] A. Philippidis, “Genexine - GX-19,” GEN, 09-Jul-2020. [Online]. Available: https://www.genengnews.com/covid-19-candidates/genexine-gx-19/. [Accessed: 05-Jan-2022].

[93] A. Dey, T. M. Chozhavel Rajanathan, H. Chandra, H. P. R. Pericherla, S. Kumar, H. S. Choonia, M. Bajpai, A. K. Singh, A. Sinha, G. Saini, P. Dalal, S. Vandriwala, M. A. Raheem, R. D. Divate, N. L. Navlani, V. Sharma, A. Parikh, S. Prasath, M. Sankar Rao, and K. Maithal, “Immunogenic potential of DNA vaccine candidate, zycov-D against SARS-COV-2 in animal models,” Vaccine, 2021.

[94] R. Ella, S. Reddy, H. Jogdand, V. Sarangi, B. Ganneru, S. Prasad, D. Das, D. Raju, U. Praturi, G. Sapkal, P. Yadav, P. Reddy, S. Verma, C. Singh, S. V. Redkar, C. S. Gillurkar, J. S. Kushwaha, S. Mohapatra, A. Bhate, S. Rai, S. Panda, P. Abraham, N. Gupta, K. Ella, B. Bhargava, and K. M. Vadrevu, “Safety and immunogenicity of an inactivated SARS-COV-2 vaccine, BBV152: Interim results from a double-blind, randomised, multicentre, phase 2 trial, and 3-month follow-up of a double-blind, Randomised Phase 1 trial,” The Lancet Infectious Diseases, vol. 21, no. 7, pp. 950–961, 2021.

[95] M. Pizza, R. Bekkat-Berkani, and R. Rappuoli, “Vaccines against meningococcal diseases,” Microorganisms, vol. 8, no. 10, p. 1521, 2020.

[96] “Understanding viral vector covid-19 vaccines,” Centers for Disease Control and Prevention. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/differentvaccines/viralvector.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fvaccines%2Fcovid-19%2Fhcp%2Fviral-vector-vaccine-basics.html. [Accessed: 04-Jan-2022].

[97] Commissioner, O.of the: “Coronavirus (COVID-19) Update: FDA Authorizes Monoclonal An-tibodies for Treatment of COVID-19”, [Online]. Available:

https://www.fda.gov/news-events/press-announce-ments/coronavirus-covid-19-update-fda-authorizes-monoclonal-antibodies-treatment-covid-19-0, last accessed 2021/5/2.

[98] Miller, K.: "Merck Oral COVID-19 Drug Shows Promise in Early Trials", [Online]. Available: https://www.verywellhealth.com/merck-oral-covid-19-drug-clinical-trial-5115909, last ac-cessed 2021/5/2.

[99] "BGB-DXP593", [Online]. Available: https://go.drugbank.com/drugs/DB16357, last accessed 2021/5/2.

[100] "A Study on the Safety, Tolerability and Immune Response of SARS-CoV-2 Sclamp (COVID-19) Vaccine in Healthy Adults" - Full Text View, [Online]. Available: https://www.clinicaltri-als.gov/ct2/show/NCT04495933.

[101] K. Dhama et al., "Plant-based vaccines and antibodies to combat COVID-19: current status and prospects", Hum. Vaccin. Immunother., vol 16, no 12, bll 2913–2920, Des 2020.

[102] "Greffex inc. – the future of genetic engineering", 11-Feb-2019. [Online]. Available at: https://www.greffex.com/,. [Toegang verkry: 05-Jan-2022].

[103] "Phase II / III Study of COVID-19 DNA Vaccine (AG0302-COVID19)" - Full Text View, [Online]. Available at: https://www.clinicaltrials.gov/ct2/show/NCT04655625, last accessed 2021/5/2.

[104] "New Data from Vaxart Oral COVID-19 Vaccine Phase I Study Suggests Broad CrossReactivity against Other Coronaviruses", [Online]. Available at: https://investors. vaxart.com/news-releases/news-release-de-tails/new-data-vaxart-oral-covid-19-vaccine-phase-i-study-suggests, last accessed 2021/5/2.

[105] N. Khan, "Review of: 'efficacy of the chadox1 ncov-19 covid-19 vaccine against the B.1.351 variant,'" 2021.

[106] “Efficacy and Safety of TY027 a Treatment for COVID-19 in Humans”, [Online]. Available at: https://www.center-watch.com/clinical-trials/listings/259289/efficacy-and-safety-of-ty027-a-treatment-for-covid-19-in-humans/, last accessed 2021/5/2.

[107] “Safety and Immunogenicity of AdCOVID in Healthy Adults (COVID-19 Vaccine Study) “- Full Text View, [Online]. Available at: https://clinicaltrials.gov/ct2/show /NCT04679909, last accessed 2021/5/4.

[108] F. Trevisan, F. Calignano, M. Lorusso, J. Pakkanen, A. Aversa, E. Ambrosio, M. Lombardi, P. Fino, and D. Manfredi, “On the Selective Laser Melting (SLM) of the AlSi10Mg Alloy: Process, Microstructure, and Mechanical Properties,” Materials, vol. 10, no. 1, p. 76, 2017.

[109] S. Liu and Y. C. Shin, “Additive manufacturing of Ti6Al4V alloy: A review,” Materials & Design, vol. 164, p. 107552, 2019.

[110] H. Baumgartl, J. Tomas, R. Buettner, and M. Merkel, “A deep learning-based model for defect detection in laser-powder bed fusion using in-situ thermographic monitoring,” Progress in Additive Manufacturing, vol. 5, no. 3, pp. 277–285, 2020.

[111] M. Kimura, A. Hirayama, J. Yoshioka, H. Maekawa, M. Kusaka, K. Kaizu, and T. Takahashi, “Mechanical Properties of AlSi12 Alloy Manufactured by Laser Powder Bed Fusion Technique,” Journal of Failure Analysis and Prevention, vol. 20, no. 6, pp. 1884–1895, 2020.

[112] V. Livescu, C. M. Knapp, G. T. Gray, R. M. Martinez, B. M. Morrow, and B. G. Ndefru, “Additively manufactured tantalum microstructures,” Materialia, vol. 1, pp. 15–24, 2018.

[113] M. Samykano, S. K. Selvamani, K. Kadirgama, W. K. Ngui, G. Kanagaraj, and K. Sudhakar, “Mechanical property of FDM printed ABS: influence of printing parameters,” The

International Journal of Advanced Manufacturing Technology, vol. 102, no. 9-12, pp. 2779–2796, 2019.

[114] P. Hackney and R. Wooldridge, “Optimisation of Additive Manufactured Sand Printed Mould Material for Aluminium Castings,” Procedia Manufacturing, vol. 11, pp. 457–465, 2017.

[115] A. Domashenkov, A. Borbély, and I. Smurov, “Structural modifications of WC/Co nanophased and conventional powders processed by selective laser melting,” Materials and Manufacturing Processes, vol. 32, no. 1, pp. 93–100, 2016.

[116] B. Zhang, N.-E. Fenineche, H. Liao, and C. Coddet, “Microstructure and Magnetic Properties of Fe–Ni Alloy Fabricated by Selective Laser Melting Fe/Ni Mixed Powders,” Journal of Materials Science & Technology, vol. 29, no. 8, pp. 757–760, 2013.

[117] W. R. Kim, G. B. Bang, J. H. Park, T. W. Lee, B.-S. Lee, S.-M. Yang, G.-H. Kim, K. Lee, and H. G. Kim, “Microstructural study on a Fe-10Cu alloy fabricated by selective laser melting for defect-free process optimization based on the energy density,” Journal of Materials Research and Technology, vol. 9, no. 6, pp. 12834–12839, 2020.

[118] A. V. Sotov, A. V. Agapovichev, V. G. Smelov, V. V. Kokareva, and M. V. Zenina, “Investigation of the Ni-Co-Cr alloy microstructure for the manufacturing of combustion chamber GTE by selective laser melting,” The International Journal of Advanced Manufacturing Technology, vol. 101, no. 9-12, pp. 3047–3053, 2018.

[119] M. Molitch-Hou, “Overview of additive manufacturing process,” Additive Manufacturing, pp. 1–38, 2018.

[120] J. Liu, T. Gottschall, K. P. Skokov, J. D. Moore, and O. Gutfleisch, “Giant magnetocaloric effect driven by structural transitions,” Nature Materials, vol. 11, no. 7, pp. 620–626, 2012.

[121] A. Aryal, S. Pandey, I. Dubenko, D. Mazumdar, S. Stadler, and N. Ali, “Magnetostructural phase transitions and large magnetic entropy changes in Ag-doped Mn1−xAgxCoGe intermetallic compounds,” MRS Communications, vol. 9, no. 1, pp. 315–320, 2019.

[122] W. Maziarz, P. Czaja, R. Chulist, A. Wójcik, Ł. Żrodowski, B. Morończyk, R. Wróblewski, and M. Kowalczyk, “Microstructure and Magnetic Properties of Selected Laser Melted Ni-Mn-Ga and Ni-Mn-Ga-Fe Powders Derived from as Melt-Spun Ribbons Precursors,” Metals, vol. 11, no. 6, p. 903, May 2021.

[123] D. A. Gabrielyan, V. V. Semenov, A. A. Uteshev, O. A. Fedyaeva, and E. G. Shubenkova, “Heat waste use for additional electricity generating using magnets thermal power plants,” Procedia Engineering, vol. 113, pp. 198–202, 2015.

[124] H. Wada, S. Tomekawa, and M. Shiga, “Magnetocaloric effect of erco2,” Journal of Magnetism and Magnetic Materials, vol. 196-197, pp. 689–690, 1999.

[125] Y. Li, H. Wang, Y. Yao, J. Xu, Z. Han, Y. Fang, L. Zhang, C. Zhang, B. Qian, and X. Jiang, “Magnetic phase diagram, magnetocaloric effect, and exchange bias in ni43mn46sn11−xgax heusler alloys,” Journal of Magnetism and Magnetic Materials, vol. 478, pp. 161–169, 2019.

[126] Y. Li, H. Wang, Y. Yao, J. Xu, Z. Han, Y. Fang, L. Zhang, C. Zhang, B. Qian, and X. Jiang, “Magnetic phase diagram, magnetocaloric effect, and exchange bias in ni43mn46sn11−xgax heusler alloys,” Journal of Magnetism and Magnetic Materials, vol. 478, pp. 161–169, 2019.

[127] A. M. Carvalho, C. S. Alves, A. de Campos, A. A. Coelho, S. Gama, F. C. Gandra, P. J. von Ranke, and N. A. Oliveira, “The magnetic and magnetocaloric properties of GD5GE2SI2 compound under hydrostatic pressure,” Journal of Applied Physics, vol. 97, no. 10, 2005.

[128] Center for Drug Evaluation and Research, “New warnings for gadolinium-based contrast agents (GBCAS) for MRI,” U.S. Food and Drug Administration. [Online]. Available: https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-fda-warns-gadolinium-based-contrast-agents-gbcas-are-retained-body. [Accessed: 10-Mar-2022].

[129] Center for Food Safety and Applied Nutrition, “Arsenic in food and dietary supplements,” U.S. Food and Drug Administration. [Online]. Available: https://www.fda.gov/food/metals-and-your-food/arsenic-food-and-dietary-supplements. [Accessed: 10-Mar-2022].

[130] M. R. Awual, M. M. Hasan, A. M. Asiri, and M. M. Rahman, “Cleaning the arsenic(v) contaminated water for safe-guarding the public health using novel composite material,” Composites Part B: Engineering, vol. 171, pp. 294–301, 2019.