|  |  |  |
| --- | --- | --- |
| <https://www.cybercoders.com/nlp-engineer-job-548188>  What You Need  Advanced degree in a quantitative discipline, e.g., Mathematics, Linguistics, Computer Science  " 2+years industry or academic experience in applied NLP a must  " Demonstrates proficiency with:  " Java and Python  " OCR libraries such as Tesseract, PyOCR, OpenCV, .NET OCR SDK, etc.  " Extracting, cleaning, preprocessing data sets. Familiarity with NumPy and Pandas  " Supervised and unsupervised machine learning techniques. This includes regression models, decision tree models, clustering, and deep learning. Hands-on experience with **Scikit-learn, Tensorflow, Keras, or PyTorch**  " Data visualization and performing model diagnostics. Understand learning curves, work with tools such as Matpoltlib, Tableau, etc.  " Familiarity with rule-based NLP, including CFG, constituency and dependency parsing, as well as their statistical variants. Experiences using NLTK, spaCy, or Stanford NLP  " Specialization in OCR is strongly preferred. Understanding of Transformers, ELMo, BERT is preferred but not required  " Experiences with healthcare industry practices and medical coding a plus, but not required    Research experience in the one of the following fields: programming languages, program synthesis, software engineering, machine learning, or human-computer interaction  *From <*[*https://careers.microsoft.com/students/us/en/job/893763/Research-Intern-Program-Synthesis-and-AI*](https://careers.microsoft.com/students/us/en/job/893763/Research-Intern-Program-Synthesis-and-AI)*>* | * + Strong toolkit of quantitative and qualitative analytical techniques rooted in business, economic and statistical analysis   + Examples of business analysis include: market competitiveness, financial analysis, social media monitoring   + Examples of statistical analysis include: linear regression, logistic regression, non-parametric statistics, probabilistic modelling, spatial modeling   + Passionate about telling stories using data   + Desire to get hands dirty working with data every day, balanced with ability to surface insights that shape new products   + Proficient with Microsoft Excel and PowerPoint, or able to generate compatible outputs using other tools, e.g. csv files   + Some experience with statistical packages (e.g. R, Python, SPSS, SAS, STATA)   + Working knowledge with Python NLP packages like **Spacy, NLTK and pycorenlp.**      * + Must be enrolled in a M.S./Ph.D. program in Computer Science or Electrical Engineering   + 3+ years of research or practical experience in applying deep learning on large scale and real world data   + Knowledge in deep learning topics including but not limited to detection, segmentation, 3D perception, and spatial-temporal analysis     *From <*[*https://www.indeed.com/viewjob?viewtype=embedded&jk=acb0f10ccefeb810&from=vjs&tk=1ei6s887j0gfc000&topwindowlocation=%2Fjobs%3Fq%3Dinternship%26l%3DWoburn%252C%2520MA%26radius%3D10%26ts%3D1599961317580%26rq%3D1%26rsIdx%3D0%26fromage%3Dlast%26newcount%3D16%26vjk%3Dacb0f10ccefeb810*](https://www.indeed.com/viewjob?viewtype=embedded&jk=acb0f10ccefeb810&from=vjs&tk=1ei6s887j0gfc000&topwindowlocation=%2Fjobs%3Fq%3Dinternship%26l%3DWoburn%252C%2520MA%26radius%3D10%26ts%3D1599961317580%26rq%3D1%26rsIdx%3D0%26fromage%3Dlast%26newcount%3D16%26vjk%3Dacb0f10ccefeb810)*>* | * + Experience with Deep Learning Techniques   + Any publications in   such as ICML, CVPR, NeurIPS  *From <*[*https://www.linkedin.com/jobs/search/?currentJobId=1909731931&distance=10&f\_LF=f\_AL&f\_TPR=r86400&geoId=102380872&location=Boston%2C%20Massachusetts%2C%20United%20States&sortBy=DD&start=350*](https://www.linkedin.com/jobs/search/?currentJobId=1909731931&distance=10&f_LF=f_AL&f_TPR=r86400&geoId=102380872&location=Boston%2C%20Massachusetts%2C%20United%20States&sortBy=DD&start=350)*>*   * Analytics / Forecasting * Artificial Intelligence / Semantic reasoning / Machine Learning * Autonomy / Collaborative swarm behaviors * Command & Control * Computational imaging * Cyber Security / Cyber Physical * Distributed Systems * Logistics * Next-generation Networking and Communications * Non-traditional computing * Photonic integrated circuits * Positioning, Navigation and Timing * Quantum Information Processing * Sensor Systems / Lidar / Radar / Undersea acoustics * Speech, Language & Multimedia * Synthetic Biology / Biotechnology   *From <*[*https://www.linkedin.com/jobs/search/?currentJobId=1918123715&distance=0&f\_TPR=r86400&geoId=104597301&keywords=research&location=Cambridge%2C%20Massachusetts%2C%20United%20States*](https://www.linkedin.com/jobs/search/?currentJobId=1918123715&distance=0&f_TPR=r86400&geoId=104597301&keywords=research&location=Cambridge%2C%20Massachusetts%2C%20United%20States)*>*  *---------------------------------------*  **—** Experience with time series analysis (Prophet, arima, deepar, lstms)  — Experience writing maintainable, testable, production-grade Python code  — Understanding of different machine learning and deep learning algorithm families and their tradeoffs  — Experience with deep learning libraries and frameworks (Scikit-learn, XGBoost, TensorFlow, Keras, PyTorch, etc.)   * + Selenium and SeleniumGrid   + Data analytics, data mining, or other data science skills   + Database experience, preferably working with Mongo databases   + Experience working with data in Information Security, Cybersecurity, or Threat Intelligence   + Experience working with bulletin boards and forums   *From <*[*https://www.indeed.com/viewjob?viewtype=embedded&jk=862e1e8a580b8334&from=vjs&tk=1eegb7vq6351p000&topwindowlocation=%2Fjobs%3Fq%3DData%2520Science%26l%3DBoston%252C%2520MA%26radius%3D10%26fromage%3D3%26vjk%3D862e1e8a580b8334*](https://www.indeed.com/viewjob?viewtype=embedded&jk=862e1e8a580b8334&from=vjs&tk=1eegb7vq6351p000&topwindowlocation=%2Fjobs%3Fq%3DData%2520Science%26l%3DBoston%252C%2520MA%26radius%3D10%26fromage%3D3%26vjk%3D862e1e8a580b8334)*>* |
| * + Active Clearance   + MS or PhD in a scientific field such as Statistics, Mathematics, Computer Science, or Data Science or 2+ years of relevant work experience   + Familiarity with handling and analyzing data at scale, for example using **Hadoop, Dask, Spark, and MapReduce**   + Working knowledge of data store tools such as SQL and Elasticsearch, and experience interacting with databases   + Experience with deep learning and neural network training, testing, and evaluation with fluency in **Tensorflow or PyTorch**   + Specialized expertise in a data-rich field such as time-series analysis, graph analytics, geospatial analysis, image processing, or Bayesian programming   *From <*[*https://www.stresearch.com/data-scientist.html*](https://www.stresearch.com/data-scientist.html)*>* | Experience managing, coding, and analyzing qualitative data using content analysis software (e.g., Dedoose, NVivo, AtlasTi). |