

BUTTE-GLENN COMMUNITY COLLEGE DISTRICT

CLASS TITLE: DATA SCIENTIST – CCC TECHNOLOGY CENTER

L221.100

BASIC FUNCTION: Under the general direction of, and in coordination with an assigned administrator, provides professional leadership for the California Community Colleges Technology Center's (CCCTC) data science and machine learning efforts. Positional emphasis will be on the application of machine learning and statistical techniques to develop and generate actionable analytics from robust algorithms. Investigates current and future technologies to ensure the CCCTC maintains the correct set of data science technologies necessary for the delivery of its services.

REPRESENTATIVE DUTIES: *(The duties recorded below are representative of the duties of the class and are not intended to cover all the duties performed by incumbent(s) of any particular position. The omission of specific statements of duties does not exclude them from the position if the scope of work is similar, related or a logical assignment to this class. The essential duties of the class are indicated with an asterisk *.)*

- *1. Drives needs assessment and uses case development to characterize opportunities for data analytics in support of the grants administered by the CCCTC.
- *2. Collaborates with Project Managers and other project team members to develop project statements, data collection, analysis and work plans.
- *3. Creates, develops, evaluates, adapts and deploys algorithms needed to carry out integrated data analysis.
- *4. Develops and maintains statistical models and algorithms to identify data patterns of interest to, and ensure applicability in, the key grants administered by the CCCTC. Uses text mining and other techniques in the building of recommender systems. Assesses the quality, viability and efficiency of analytical models.
- *5. Refines algorithms through the application of machine learning to provide predictive models. Applies these models to the area of learning analytics. Articulates and documents the models to internal and external stakeholders.
- *6. Conducts applied research on novel methods for information extraction, representation and analysis targeted at the interpretation of application use patterns, achievement, course and discipline selection choices and financial data.
- *7. In collaboration with data science and research staff, publishes research results in peer-reviewed journals and conferences.
- *8. Develops novel techniques for pattern recognition, machine learning, predictive analytics and data sampling problems.
- *9. Utilizes code written in languages, to include, but not limited to, R, Java, Python and/or statistical packages/languages such as SPSS for modeling, optimization, simulation and statistical data analysis.
- *10. Works with other technical team members to assist in the implementation of models at scale.
- *11. Provides advanced support for projects, the implementation of project work plans and in the preparation of project deliverables.
- *12. Participates in the documentation of project writing and in the preparation of materials needed for project meetings.
- *13. Participates in the preparation of mandated and required reports.
- *14. Actively contributes to the design of project-specific data collection efforts including interviews, focus groups and surveys.
- *15. Prepares tables, graphs, fact sheets and written reports summarizing research results. Uses excellent communication skills to promote new methodologies and to present research results. Prepares oral and written reports and presentations of all project findings, as required.
16. Supports ad hoc data requests and reporting to meet the needs of the CCCTC.
17. May serve on a variety of District committees as requested.
18. Performs related work as required.

MINIMUM QUALIFICATIONS:

- **EDUCATION AND EXPERIENCE:**
 - Master's Degree in informatics, statistics, computer science or social science, or a related field with an emphasis in quantitative analysis and/or machine learning; **AND**
 - Three (3) years of experience in institutional research, data analysis, strategic or academic planning, business intelligence, decision-support, data mining, or related fields.
- **CERTIFICATES, LICENSES, REGISTRATION AND OTHER REQUIREMENTS:**
 - Hold and maintain a valid driver's license throughout duration of employment with the District.
 - Travel is required.

DESIRED QUALIFICATIONS:

- Experience with machine learning techniques, especially in the development of scalable algorithms.
- Experience with cloud infrastructure in support of data science.
- Experience with technology related to higher education.
- Experience with statistical analysis software and techniques.
- Experience with programming and leveraging a language.
- Experience in an institutional research position.

KNOWLEDGE SKILLS AND ABILITIES: (May be acquired through education, training and/or experience.)

Knowledge of: District policies and procedures and rules and regulations; laws and ordinances pertaining to the specific responsibilities of the position; managerial concepts and principles; educational systems, specifically the California Community College system; significant business strategy skills; a broad level of familiarity with statistical techniques and predictive analytics; standard machine learning, data and web mining techniques; transforming ideas into solutions; natural language processing; predictive modeling; PPML and its application; graph analysis; error propagation; inventing techniques; model fit measures; confidence intervals; structural equivalence models; (dis)similarity measures such as Jaccard Index.

Ability to: ensure departmental compliance with all laws, regulations and policies and procedures of the District and those related to the position; motivate and provide successful leadership; comprehend and stay current with evolving technologies within multiple areas; effectively implement project management methods to ensure increased productivity; utilize multiple programming languages to build and deploy analytical models; code with object oriented language; design and develop software products; use statistical modeling tools such as Rapid Miner, Stata, SPSS, KNIME, Orange Canvas or SAS and extend them with R (R Studio), Java (WEKA) or Python code; utilize effective machine learning techniques; utilize tools such as Map Reduce and/or Hadoop (MapR, HortonWorks, etc); develop system prototypes; communicate findings; extract and query large data sets; refine applicable parameters; accurately express data trends and/or anomalies; exhibit the understanding of predictive analytics; effectively multi-task and coordinate multiple activities simultaneously; exhibit the ability to analyze data and utilize strong problem solving skills; evaluate various algorithmic approaches; accurately assess the technological needs of a department or program; utilize keyboarding skills commensurate with the required functions for this position; compose complex, technical reports, analysis and correspondence; analyze administrative problems and prepare comprehensive, clear and concise reports and recommendations; make sound, educated decisions; work independently; establish and maintain cooperative working relationships with those contacted in the course of work; work with diverse populations; make required mathematical calculations accurately; maintain records accurately and with confidentiality as required; compile, organize and use various financial information necessary in the preparation of the departmental budget; control and monitor budget expenditures and resource development.

WORK DIRECTION, LEAD AND SUPERVISORY RESPONSIBILITIES:

May provide direct supervision to professional and classified staff; may provide supervision to student employees and interns; may provide technical direction to faculty members.

PHYSICAL EFFORT:

Normal office environment. Travel is required.

CONTACTS:

Faculty, staff, administrators, students and the general public.

WORKING CONDITIONS:

Normal office environment. Subject to interruptions, noise and computer-generated video radiation.

NOTE: THIS CLASS IS EXEMPT UNDER FLSA PROVISIONS

Butte-Glenn Community College District is an Equal Opportunity Employer. In compliance with the Americans with Disabilities Act, Butte-Glenn Community College District will provide reasonable accommodation to qualified individuals. Butte-Glenn Community College District encourages both incumbents and individuals who have been offered employment to discuss potential accommodations with the employer.