

معاونت پژوهش، برنامه ریزی و سنجش مهارت
دفتر پژوهش، طرح و برنامه ریزی درسی

استاندارد شایستگی

تحلیل داده با پایتون

(با رویکرد علم داده)

گروه شغلی

فناوری اطلاعات

کد ملی استاندارد

۲	۵	۲	۱	۴	۰	۵	۳	۰	۰	۰	۰	۰	۱	۱
ISCO-08				سطح مهارت	شناسه گروه			شناسه شغل			شناسه شایستگی		نسخه	

تاریخ تدوین استاندارد: ۱۴۰۱/۸/۱

شناسنامه استاندارد شایستگی

استاندارد شایستگی تحلیل داده با پایتون (با رویکرد علم داده) با کد ۲۵۲۱۴۰۵۳۰۰۰۰۱۱ با مشارکت خبرگان حرفه‌ای، صاحبان مشاغل، مربیان و کارشناسان برنامه‌ریزی درسی تدوین و در جلسه مورخ ۱۴۰۱/۰۷/۱۷ گروه شغلی فناوری اطلاعات بررسی و به تصویب رسید و در سامانه ملی استاندارد مهارت بارگذاری گردید.

کلیه حقوق مادی و معنوی این استاندارد متعلق به سازمان آموزش فنی و حرفه‌ای کشور بوده و هرگونه سوء استفاده مادی و معنوی از آن موجب پیگرد قانونی است.
از صاحبان نظران، متخصصان، صاحبان مشاغل و خبرگان حرفه‌ای دعوت بعمل می‌آید با ارسال نقطه نظرات اصلاحی خود، در انطباق بیشتر این استاندارد با نیازهای اکنون و آینده بازار کار، مشارکت نمایند.

سازمان آموزش فنی و حرفه‌ای کشور

معاونت پژوهش، برنامه‌ریزی و سنجش مهارت

دفتر پژوهش، طرح و برنامه‌ریزی درسی

تهران، خیابان آزادی، نبش خیابان خوش جنوبی

تلفن: ۶۶۵۸۳۶۲۸ - ۰۲۱

ایمیل: rpc@irantvto.ir

سامانه ملی استاندارد مهارت: rpc.irantvto.ir



نظارت بر تدوین محتوا و تصویب استاندارد: دفتر پژوهش، طرح و برنامه ریزی درسی

کد ملی شناسایی استاندارد شایستگی: ۲۵۲۱۴۰۵۳۰۰۰۰۰۱۱

اعضاء کارگروه برنامه‌ریزی درسی: فناوری اطلاعات						
ردیف	نام و نام خانوادگی	آخرین مدرک تحصیلی	رشته تخصصی	شغل	سمت در کارگروه	سابقه/تجربه کار
۱	محبوبه بهروز	کارشناسی ارشد	مهندسی کامپیوتر- نرم افزار	مربی فنی و حرفه ای و مدیرعامل شرکت بهین برازش رایان پویش	شاغل خبره بازار کار	۳
۲	علی ملائی	دکترا	ارشد هوش مصنوعی و دکترای مدیریت راهبردی امنیت فضای سایبر	استاد دانشگاه و مشاور فنی شرکت بهین برازش رایان پویش	شاغل خبره بازار کار	۱۵
۳	جواد ایزدی	کارشناسی ارشد	مهندسی کامپیوتر- هوش مصنوعی	کارشناس پایگاه داده و تحلیلگر داده	شاغل خبره بازار کار	۱۲
۴	احمد حنتی فرد	دکترا	مهندسی کامپیوتر- هوش مصنوعی	مربی فنی و حرفه ای- پژوهشگر حوزه هوش مصنوعی	شاغل خبره بازار کار	۱۵
۸	محمد رضا کنجه مرادی	کارشناسی ارشد	مدیریت IT	کارشناس طرح و برنامه ریزی درسی گروه فناوری اطلاعات	کارشناس طرح و برنامه ریزی درسی گروه فناوری اطلاعات	۱۲
۹	محمد رضا اسدی	کارشناسی ارشد	مهندسی IT	دبیر کارگروه برنامه ریزی درسی گروه فناوری اطلاعات	دبیر کارگروه برنامه ریزی درسی گروه فناوری اطلاعات	۲۵

تعاریف اصطلاحات بکار رفته در این استاندارد

شایستگی: توانایی انجام کار در محیطها و شرایط گوناگون به طور موثر و کارا برابر استاندارد می‌باشد.

کد: مجموعه‌ای از اعداد برگرفته از سند بین المللی طبقه بندی مشاغل ISCO-2008 و شناسه‌های قراردادی است که به منظور شناسایی استانداردها تعیین می‌شود.

حرفه: مجموعه‌ای از چند شغل همگن که در یک گروه بزرگ دسته‌بندی شده، در مبانی و اصول مشترک بوده و در طبقه‌بندی شغل و حرف در یک سطح مهارت و تخصص قرار می‌گیرند.

استاندارد شایستگی: نقشه کاری شامل مولفه‌های شایستگی، استاندارد عملکرد و تجهیزات و ابزار مورد نیاز برای انجام کار در شغل و حرفه مورد نظر می‌باشد.

شرح استاندارد: بیانیه‌ای شامل مهم‌ترین عناصر یک شغل یا شایستگی از قبیل جایگاه، کارها، ارتباط با مشاغل دیگر در یک حوزه شغلی، مسئولیت‌ها، شرایط کاری و استاندارد عملکرد مورد نیاز می‌باشد.

مراحل کار: مراحل کار از تجزیه و تحلیل حرفه، وظایف و تکالیف کاری نشأت گرفته و نشان دهنده مرحله یک کار عملی بوده و قابل اندازه گیری و مشاهده می‌باشد.

استاندارد عملکرد: سطح قابل قبول برای انجام یک کار در یک حرفه یا شغل را استاندارد عملکرد کار می‌نامند. که شامل انجام کار با استفاده از تجهیزات، ابزار و مطابق با استانداردهای کاری ملی و بین المللی می‌باشد.

اعتبار: تعیین کننده ارزش آموزشی هر شایستگی است و هر واحد آن معادل ۳۰ ساعت آموزش می‌باشد.

سطح مهارت: نشان دهنده سطح مهارت مطابق ۴ سطح سند ISCO-2008 است.

سطح ایسکد مورد نیاز: نشان دهنده حداقل سطح آموزش مورد نیاز برای کار در سطح مهارت خاص مطابق با سند ISCO-2008 می‌باشد.

مولفه‌های شایستگی: شامل اجزای شایستگی دانش، مهارت و نگرش مورد نیاز برای انجام کار در شغل و حرفه مورد نظر می‌باشد.

دانش: حداقل مجموعه‌ای از معلومات نظری و توانمندی‌های ذهنی لازم برای رسیدن به یک شایستگی است که می‌تواند شامل علوم پایه، تکنولوژی و زبان فنی باشد.

مهارت: حداقل هماهنگی بین ذهن و جسم برای رسیدن به یک توانمندی یا شایستگی است که معمولاً به مهارت‌های عملی ارجاع می‌شود.

نگرش: مجموعه‌ای از رفتارهای عاطفی که برای شایستگی در یک کار مورد نیاز است و شامل مهارت‌های غیرفنی و اخلاق حرفه‌ای می‌باشد.

ایمنی و بهداشت: مواردی است که عدم یا انجام ندادن صحیح آن موجب بروز حوادث و خطرات در محیط کار می‌شود.

توجهات زیست محیطی: ملاحظاتی است که در هر شغل باید رعایت شود تا کمترین آسیب به محیط زیست وارد گردد.

تجهیزات، ابزار و مواد مصرفی: حداقل امکانات مورد نیاز برای اجرای آموزش مطابق استاندارد شایستگی است.

کد: ۲۵۲۱	عنوان حرفه: طراحان و مدیران پایگاه داده ها
کد: ۲۵۲۱۴۰۵۳۰۰۰۰۰۱۱	عنوان استاندارد شایستگی: تحلیل داده با پایتون (با رویکرد علم داده)

شرح استاندارد

این استاندارد شامل دانش، مهارت و نگرش لازم جهت انجام کار « تحلیل دیتاهای ساختاریافته و بدون ساختار از پایگاه داده های سازمانها و شرکتهای » در محیطها و شرایط گوناگون به طور موثر و کارا بر اساس استانداردهای محیط کار و مراحل کاری زیر می باشد.

مراحل کار

۱. نصب ابزارهای مورد نیاز علم داده
۲. روش شناسی علم داده
۳. برنامه نویسی پایتون برای علم داده و هوش مصنوعی
۴. تحلیل داده با پایتون
۵. مصورسازی داده با پایتون
۶. یادگیری ماشین با پایتون
۷. پردازش تصویر مقدماتی

استاندارد عملکرد

تحلیل داده با پایتون (با رویکرد علم داده) با استفاده از ابزار و کتابخانه های زبان برنامه نویسی پایتون و بر اساس مستندات زبان برنامه نویسی پایتون و راهنمای نرم افزار Anaconda

سطح مهارت شایستگی

سطح چهار (۴) مهارت

سطح ایسکد مورد نیاز

کارشناسی

اعتبار

دو (۲) واحد

ردیف	مراحل کار	مولفه های شایستگی
۱	نصب ابزارهای مورد نیاز علم داده	<p>دانش:</p> <ul style="list-style-type: none"> - مبانی علم داده - مسیرهای رسیدن به علم داده - مفاهیم هوش مصنوعی مرتبط با علم داده - علم داده - نوت بوک ژوپیتتر (Jupyter Notebook) - فایل ها و Package ها و کتابخانه های پایتون <p>مهارت:</p> <ul style="list-style-type: none"> - نصب نرم افزار Anaconda - کار با محیط نرم افزار Jupyter Notebook - بکارگیری دستورات pip جهت نصب پکیج های مورد نیاز - راه اندازی محیط آناکاندا و کد نویسی
۲	روش شناسی علم داده	<p>دانش:</p> <ul style="list-style-type: none"> - مدل کسب و کار و شناسایی کردن مسئله - شناسایی شاخص های کلیدی (KPIs) کسب و کار و الزامات جمع آوری داده - روش طرح فرضیه و مدل سازی - نحوه ارزشیابی مدل و ارائه نتایج در غالب نمودارها و داده داستان (Data Story Telling) <p>مهارت:</p> <ul style="list-style-type: none"> - گردآوری، تجزیه و تحلیل کردن داده های محیط کسب و کار - تعیین کردن متغیرها و ویژگی های مهم و تاثیرگذار بر متغیر وابسته ی مسئله - پیدا کردن مدل مناسب، شکل دهی فرضیه و راهکارهای حل مسئله برای مدل سازی
۳	برنامه نویسی پایتون برای علم داده و هوش مصنوعی	<p>دانش:</p> <ul style="list-style-type: none"> - زبان برنامه نویسی پایتون - عبارات و انواع متغیرها - انواع عملگرهای منطقی و مقایسه ای - انواع داده در پایتون

مولفه های شایستگی	مراحل کار	ردیف
<ul style="list-style-type: none"> - ساختار داده در پایتون - نوع های داده لیست، تاپل، دیکشنری، مجموعه - نحوه برنامه نویسی پایتون - شرط ها، حلقه ها، توابع، شی و کلاس - نحوه خواندن و نوشتن فایل - کتابخانه های مهم علم داده در پایتون (numpy, pandas, sklearn, scipy,...) 		
<p>مهارت:</p> <ul style="list-style-type: none"> - کدنویسی با زبان برنامه نویسی پایتون و استفاده کردن از انواع داده - کار با کتابخانه های مهم علم داده در پایتون (numpy, pandas, sklearn,) (scipy,...) - پیاده سازی مدل ها در غالب برنامه های کاربردی 		
<p>دانش:</p> <ul style="list-style-type: none"> - فهم داده، ورود و صدور داده در پایتون، تفکیک داده، مواجهه با داده از دست رفته - نحوه شکل دهی، نرمال کردن داده، تبدیل داده های طبقه بندی به داده های کمی، تحلیل توصیفی داده، همبستگی، P_value و تحلیل واریانس - مدل های دسته بندی، رگرسیون، مصورسازی، ارزیابی مدل و مدلسازی - شرایط کارایی مدل (accuracy, recall, presicion, f-measure) - نحوه معرفی overfitting, underfitting 	تحلیل داده با پایتون	۴
<p>مهارت:</p> <ul style="list-style-type: none"> - تمیز کردن داده، استاندارد سازی و نرمال سازی داده ها، بررسی کردن داده های از دست رفته و انجام دادن کلیه مراحل پیش پردازش داده برای مدلسازی - تحلیل توصیفی داده ها - مدلسازی و مصورسازی داده - ارزیابی و کارایی مدل پیاده سازی شده 		

مولفه های شایستگی	مراحل کار	ردیف
<p>دانش:</p> <ul style="list-style-type: none"> - روش معرفی داده - کتابخانه های مصورسازی: (seaborn, folium Matplotlib, ...) - نمودارهای Box, Charts Pie, Bar Charts, Histograms, Area Plots, Plots, Word Clouds, Scatter Plots <p>مهارت:</p> <ul style="list-style-type: none"> - کار با کتابخانه های مصورسازی در پایتون (matplotlib, seaborn, ...) و رسم انواع نمودارهای تحلیلی - تحلیل و توصیف داده از روی نمودار - استخراج دانش و نمایش داده ها در قالب نمودارها 	<p>مصورسازی داده با پایتون</p>	<p>۵</p>
<p>دانش:</p> <ul style="list-style-type: none"> - انواع یادگیری (با نظارت، بدون نظارت و نیمه نظارت شده supervised, (unsupervised, semi-supervised learning) - الگوریتم های یادگیری ماشین - Evaluation Metrics in Classification, K-Nearest Neighbours - Intro to Building Decision Trees, Introduction to Decision Trees - Logistic regression vs Linear regression, Logistic Regression - Intro to Support Vector Machine, Logistic Regression Training - Intro to More on k-Means, Intro to k-Means, Clustering - More on Hierarchical Clustering, Hierarchical Clustering - Collaborative, Systems, Content-based Recommender, DBSCAN, Filtering 	<p>یادگیری ماشین با پایتون</p>	<p>۶</p>
<p>مهارت:</p> <ul style="list-style-type: none"> - پیاده سازی الگوریتم های یادگیری ماشین - بهینه سازی های پارامترهای الگوریتم های یادگیری ماشین - بکارگیری هر مدل متناسب با نوع داده 		

مولفه های شایستگی	مراحل کار	ردیف
<p>دانش:</p> <ul style="list-style-type: none"> - روش کتابخانه OpenCV، نوع داده تصویر، معرفی فضاهای رنگ - روش عملیات روی تصاویر، انواع اعمال فیلترها بر تصاویر، هیستوگرام تصاویر، استخراج خطوط و مرزها، ایجاد کانتورها - علائم راهنمایی با وب کم و پردازش ویدئو، دسته بندی تصاویر اعداد دست نوشته و دسته بندی تصاویر - روش تطبیق الگو - تبدیلات مورفولوژیک 	پردازش تصویر مقدماتی	۷
<p>مهارت:</p> <ul style="list-style-type: none"> - پردازش داده تصویر - کار با انواع فیلترهای تصویری - کار با کتابخانه opencv - کار با الگوریتم های یادگیری ماشین و پیش پردازش داده های تصویر - پیاده سازی یک شبکه عصبی نمونه برای دسته بندی تصاویر 		

نگرش

- رعایت دقت و نظم در انجام کار
- رعایت اخلاق حرفه ای
- بکارگیری خلاقیت و کارآفرینی

ایمنی و بهداشت

- رعایت ارگونومی هنگام کار
- رعایت بهداشت فردی و عمومی در محیط کار

توجهات زیست محیطی

- تفکیک صحیح زباله و قطعات مستهلک شده کامپیوتر

تجهیزات		
مشخصات فنی	عنوان	ردیف
استاندارد	پرده نمایش یا بورد هوشمند	۱
استاندارد	تخته وایت بورد	۲
استاندارد	دیتا پروژکتور با متعلقات	۳
حداقل Corei5 - حداقل ۸ گیگا بایت رم - مانیتور LED	رایانه با متعلقات	۴
CO2	کپسول اطفاء حریق	۵
چرخدار - چرمی	صندلی مربی	۶
دسته دار	صندلی کارآموز	۷
	تجهیزات اتصال به اینترنت	۸

ابزار		
مشخصات فنی	عنوان	ردیف
آخرین ورژن	نرم افزار Anaconda	۱
ویندوز ۸ به بالا	سیستم عامل	۲

مواد مصرفی		
مشخصات فنی	عنوان	ردیف
در ۴ رنگ ۲ سری	ماژیک وایت برد	۱
معمولی (به تعداد لازم)	CD یا DVD خام	۲
معمولی (به تعداد لازم)	فلش مموری	۳